

The Role of MIER1 and G9a in Histone 3 Lysine 27 (H3K27) Methylation

By Leena Derwish

A thesis submitted to the
School of Graduate Studies
in partial fulfillment of the
requirements for the degree of
Masters in Science.

Division of Biomedical Sciences
Faculty of Medicine
Memorial University
October, 2014

Abstract

G9a is a lysine methyltransferase, which is responsible for mono- and dimethylation of lysine 9 on histone 3 in euchromatic loci containing transcriptionally active genes. G9a has been shown to be implicated in the methylation of lysine 27 on histone 3 as well. This study focused on characterizing the interaction between G9a and mesoderm induction early response 1 (MIER1), a fibroblast growth factor (FGF)-activated transcriptional regulator. Glutathione S-transferase pull-down assay showed that MIER1 β interacts with G9a via the SANT domain, a domain common in MIER1 isoforms. Co-immunoprecipitations using lysates from human embryonic kidney cells (HEK-293) confirmed the validity of the interaction between these two proteins *in vivo*. Co-immunoprecipitations also revealed MIER1 α to interact with chromodomain protein, Y-like (CDYL). Moreover, CDYL bridges enhancer of zeste homolog 2 (EZH2), the methyltransferase subunit of PRC2 complex, to MIER1 α . Methyltransferase assays using histone H3 N-terminal [H3(21-44)] peptide or recombinant histone H3 monomer as substrates showed recombinant G9a to have no detectable methylation activity towards H3K27. GST-MIER1 β does not have intrinsic histone methyltransferase activity and does not methylate H3K27. However, the MIER1 α complex immunoprecipitated from HEK-293 cells does methylate H3K27. A methyltransferase associated with MIER1 α in HEK-293 cells, such as G9a or EZH2, could be responsible for this methylation.

Acknowledgements

First and foremost, I would like to take this opportunity to extend an earnest thank you to my supervisor, Dr. Gary Paterno, for his unwavering, unyielding support and encouragement. I would like to further thank him for helping me better understand the theories and principles relating to my project. I would also like to thank my supervisory committee, Dr. Laura Gillespie and Dr. Robert Gendron, for keeping me focused on the essential questions and for their valuable counselling throughout the duration of my project.

I express my utmost gratitude to Corinne Mercer for her exceptional technical support. I would not have been able to complete my master's project without her guidance in mastering various techniques and protocols. I would also like to acknowledge all the members of the Terry Fox Cancer Research Labs – Aimee, Jarratt, Julia, Phillip, Roya, Satoko, Shengnan, Tahrim, and Youlian – for always being there, helping me when I needed assistance, and most importantly, for their friendship. The past three years working at the Terry Fox Labs have been most rewarding and I will cherish each memory forever.

Funding for this project was provided in part by the Canadian Institutes of Health Research and Memorial University of Newfoundland.

Table of Contents

Abstract	ii
Acknowledgements	iii
Table of Contents	iv
List of Tables	vii
List of Figures	viii
Chapter 1: Introduction	1
1.1 General introduction.....	1
1.1.1 Epigenetics	1
1.1.2 DNA methylation	4
1.1.3 Chromatin structure and histone modifications	4
1.1.4 Histone code hypothesis.....	5
1.1.5 Writers, erasers and readers of the histone code	5
1.1.6 Histone lysine methylation.....	7
1.1.7 H3K27 methylation.....	7
1.1.8 Histone lysine methyltransferases (HKMTs).....	9
1.1.9 Recruitment of histone methyltransferases to their genomic destination.....	12
1.2 Euchromatic human methyltransferase 2 (G9a)	14
1.2.1 G9a	14
1.2.3 G9a structure	14
1.2.4 The effect of other post-translational modifications on G9a activity	18
1.2.5 Recruiting other proteins to reinforce silencing	18
1.2.6 The role of G9a: biological relevance	20
1.2.7 G9a inhibitors.....	22
1.3 Chromodomain, Y-like: CDYL.....	25
1.31 CDYL.....	25

1.3.2	CDYL domains and function	25
1.4	<i>Mesoderm induction early response 1: MIER1</i>	28
1.4.1	MIER1	28
1.4.2	MIER1 structure and function.....	28
1.5	Hypothesis and objectives.....	33
Chapter 2: Materials and Methods		37
2.1	<i>Plasmids</i>	37
A.	GST-MIER1 β , GST-aa1-283, GST-aa287-512, GST-aa287-357.....	37
B.	GST-SANT mutant.....	37
C.	pCS3+MT, myc-MIER1 α	37
D.	myc-G9a-SET	38
E.	myc-G9a, myc-G9a-N-terminal, myc-G9a-Ank.....	47
F.	pCMV-Tag2B, Flag-CDYL.....	49
2.11	Plasmid preparation.....	52
2.2	<i>Cell culture</i>	54
2.3	<i>Transfection of HEK-293 cells</i>	54
2.4	<i>GST pull-down assay</i>	55
2.4.1	GST-fusion protein production	55
2.4.2	GST pull-down assay	57
2.5	<i>Immunoprecipitation</i>	59
2.5.1	Cell lysate preparation	59
2.5.2	Co-immunoprecipitation	60
2.5.3	Western blot analysis	61
2.6	<i>Methylation assay using recombinant G9a protein</i>	62
2.6.1	Methylation assay using H3(21-44) peptide as the substrate	62
2.6.2	Methylation assay using recombinant histone 3 as the substrate	64
2.7	Methylation assay using GST-MIER1 β protein and recombinant H3(21-44) peptide as the substrate	66
2.8	Methylation assay using MIER1 α complex from HEK293 cells	67
Chapter 3: Results		72

3.1	<i>Interaction between MIER1 and G9a</i>	72
3.1.1	<i>In vitro</i> interaction between MIER1 β and G9a	72
3.1.2	<i>In vivo</i> interaction between MIER1 α and G9a	75
3.2	<i>Interaction between MIER1α, G9a, CDYL and EZH2</i>	78
3.2.1	<i>In vivo</i> interaction between MIER1 α , G9a and CDYL	78
3.2.2	<i>In vivo</i> interaction between G9a and CDYL	81
3.2.3	<i>In vivo</i> interaction between MIER1 α , CDYL and EZH2	84
3.3	<i>Histone 3 lysine 27 methylation using recombinant G9a protein</i>	87
3.3.1	Methylation assay using H3(21-44) peptide as the substrate	87
3.3.2	Methylation assay using recombinant histone as the substrate	89
3.3.3	Methylation assay using recombinant histone 3 (H3) as the substrate: effects of different inhibitors on G9a activity	91
3.4	<i>Histone 3 lysine 27 methylation using GST-MIER1β</i>	93
3.5	<i>Histone 3 lysine 27 methylation by endogenous MIER1α complex</i>	95
Chapter 4: Discussion		98
4.1	<i>General discussion</i>	98
4.2	<i>Proposed role of MIER1 in H3K27 methylation</i>	110
4.3	<i>Implications</i>	114
4.4	<i>Future studies</i>	115
4.4.1	Determine <i>in vivo</i> interaction between endogenous MIER1 α , G9a and CDYL	115
4.4.2	Demonstrate <i>in vitro</i> interaction between MIER1 α and CDYL	116
4.4.3	Determine whether MIER1 interacts with other members of PRC2 complex	116
4.4.4	Determine the methylation activity of full length G9a	116
4.4.5	Determine the methylation activity of MIER1 complex	117
Bibliography		118

List of Tables

Table 1: PCR primers.....	40
Table 2: Reagents for PCR insert amplification	41
Table 3: PCR setting for PCR cloning of the insert.....	42
Table 4: Reagents for PCR verification of the insert	43
Table 5: PCR setting for verification of the insert	44
Table 6: Plasmids	53
Table 7: Primary antibodies	70
Table 8: Secondary antibodies	71

List of Figures

Figure 1: Waddington's illustrations.....	3
Figure 2: S-adenosyl-methionine is the methyl donor in methylation of histone lysine residues	11
Figure 3: Domain structure of G9a	17
Figure 4: Domain structure of MIER1	32
Figure 5: Hypothesised role of MIER1 in H3K27 methylation.....	36
Figure 6: GST pull-down showing the interaction between G9a and GST-MIER1 β fusion constructs <i>in vitro</i>	74
Figure 7: MIER1 α interacts with G9a in HEK 293 cells.....	77
Figure 8: MIER1 α interacts with G9a and CDYL in HEK 293 cells.	80
Figure 9: G9a and CDYL do not interact <i>in vivo</i> in HEK 293 cells	83
Figure 10: CDYL bridges EZH2 to MIER1 <i>in vivo</i>	86
Figure 11: Recombinant G9a-SET protein weakly methylates lysine 27 on histone H3 peptide.....	88
Figure 12: Recombinant G9a-SET protein methylates lysine 9 but not 27 on histone H3	90
Figure 13: The activity of recombinant G9a-SET protein is inhibited by S-adenosyl homocysteine (SAH), and the G9a specific inhibitor, BIX-01294, but not GSK-343.	92
Figure 14: GST-MIER1 β does not methylate lysine 27 on histone H3 peptide	94
Figure 15: MIER1 α complex methylates lysine 27 on H3(21-44) peptide	97
Figure 16: Proposed role of MIER1 in H3K27 methylation.....	113

List of Abbreviations

°C	degrees Celsius
µg	microgram
µl	microlitre
µM	micromolar
5-mdC	5-methyl-2'-deoxycytidine
A	ampere
aa	amino acid
ACINUS	apoptotic chromatin condensation inducer in the nucleus
BSA	bovine serum albumin
CBP	CREB-binding protein
CLBC	Claudine-low breast cancer
CTH	calf thymus total histones
C/EBP-β	CCAAT/Enhancer-binding protein-β
CDY	human chromodomain Y protein
CDYL	human chromodomain, Y-like protein
Cdyl	mouse chromodomain, Y-like protein
CpG	cytosine phosphate guanine
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
DNMT	deoxyribonucleic acid methyltransferase(s)
DOT1	disruptor of telomeric silencing 1
DOT1L	disruptor of telomeric silencing-1-like
EDTA	ethyldiamine tetra-acetic acid
ELM2	EGL27 and MTAI homology 2
EHMT1	euchromatic human methyltransferase 1, also known as G9a-like protein
EHMT2	euchromatic human methyltransferase 2, also known as G9a
Ep-CAM	epithelial cell adhesion molecule
ESC	extra sex combs
EZH1/2	enhancer of zeste homologue 1/2
EED	embryonic ectoderm development
ERα	estrogen receptor α
GLP	g9a-like protein
GST	glutathione S-transferase
GST-M1ERI	glutathione S-transferase tagged MIER1
H1	histone 1
H2A	histone 2A
H2B	histone 2B
H3	histone 3
H3K9	histone 3 lysine 9
H3K9me1	mono-methylated histone 3 lysine 9

H3K9me2	di-methylated histone 3 lysine 9
H3K27	histone 3 lysine 27
H3K27ac	acetylated histone 3 lysine 27
H3K27me1	mono-methylated histone 3 lysine 27
H3K27me2	di-methylated histone 3 lysine 27
H3K27me3	tri-methylated histone 3 lysine 27
H4	histone 4
HAT	histone acetyltransferase(s)
HDAC	histone deacetylase(s)
HKMT	histone lysine methyltransferase(s)
HKDM	histone lysine demethylase(s)
HMG1	high mobility group 1 protein
HP1	histone protein 1
Hox	homeobox
HOTAIR	homeobox transcript antisense RNA
hr	hour(s)
HRP	horseradish peroxidase
MIER1	mesoderm induction early response 1
Igλ L	immunoglobulin lambda light chain
IL-3	interleukin 3
IL-4	interleukin 4
IL-5	interleukin 5
IP	immunoprecipitation
K_d	dissociation constant
kDa	kilodalton
KDM1A	lysine demethylase 1A
KMT	lysine methyltransferase
L3MBTL1	lethal(3)malignant brain tumor-like protein
LPS	lipopolysaccharide
lncRNAs	long non-coding RNAs
LXXLL	L= leucine, X= any amino acid; nuclear receptor binding motif
M	molar
MBD	methyl-CpG-binding domain
MBT	malignant brain tumour
min	minute(s)
mESCs	murine embryonic stem cells
ml	millilitre(s)
mM	millimolar
mRNA	messenger ribonucleic acid
NAc	nucleus accumbens
N-CoR	nuclear receptor co-repressor 1
NES	nuclear export signal
ng	nanogram

nM	nanomolar
NLS	nuclear localization signal
NRSF	neuron-restrictive silencing factor
PBS	phosphate buffered saline
PcG	polycomb group
PRC	polymerase chain reaction
PHD	plant homeodomain
PHORC	pleiohomeotic repressive complex
PMSF	phenylmethanesulfonyl fluoride
PRC1	polycomb repressive complex 1
PRC2	polycomb repressive complex 2
PREs	polycomb group response elements
PRDI-BF1	positive regulatory domain I-binding factor 1
PRMT	protein arginine methyltransferase(s)
PTM	post translational modification
RbAp46/48	retinoblastoma associated protein 46/48
REST	RE1-binding silencer of transcription
RNA	ribonucleic acid
RNAi	ribonucleic acid interference
SAH	s-adenosyl-homocysteine
SAM	s-adenosyl-methionine
SANT	domain first identified in SWI3, ADAI, NCoR, TFIIIB
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate- polyacrylamide gel electrophoresis
SET	SU(VAR)3-9, enhancer-of-Zeste, Trihorax domain
SMYD3	SET and MYND domain-containing protein 3
S _N 2	nucleophilic substitution reaction
SRA	SET- or RING-associated protein
SUV39H1/2	suppressor of variegation 3-9 homolog 1/2
SUZ12	suppressor of zeste 12
TBST	tris-buffered saline-tween 20
Th2	type 2 helper T cells
TREs	trithorax group response elements
TrxG	trithorax group
V	volt
WD40-repeat	W: tryptophan, D; aspartic acid
WDR5	WD repeat domain 5; W: tryptophan, D; aspartic acid
WIZ	widely interspaced zinc fingers

Chapter 1: Introduction

1.1 General introduction

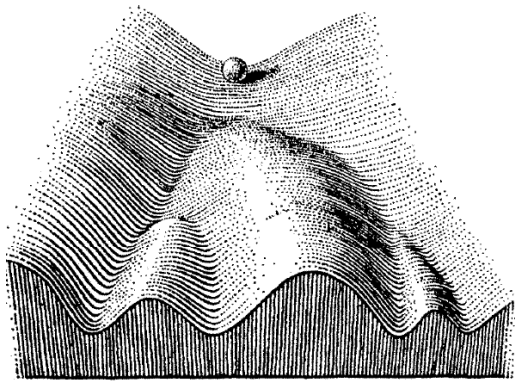
1.1.1 Epigenetics

Epigenetics is a term first coined by the British embryologist Conrad Waddington in 1942 in the context of developmental biology. Waddington believed that the interactions between networks of genes play an important role in embryonic development. His view on development, which he illustrated in “epigenetic landscapes” (Fig. 1), was what we now would call differential gene expression and regulation. In his illustrations, the developmental system is depicted as a landscape where bifurcating and deepening valleys run down from a plateau [1, 2]. In Fig. 1A, the fertilized egg is represented by the rippling plateau and the direction the ball rolls signifies the developmental path from the egg to a particular tissue or organ at the end of a valley. Moreover, the paths, slopes, and the cross-sections of the valleys are under genetic control, which is represented in Fig. 1B; the pegs represent the genes and the figurative ropes represent the “chemical tendencies” of the gene. Using this figure, Waddington attempted to illustrate that the relationship between a gene and its phenotypic effects is not simple because “if any gene mutates, altering the tension in a certain set of guy ropes, the result will not depend on that gene alone, but on its interactions with all the other guys”[2]. Overall, Waddington used the term epigenetics to describe “the branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being”[3].

However, with increased knowledge of chromatin biology, epigenetics has now been adapted to a more chromatin centric meaning.

Epigenetics is currently defined as an alteration in gene expression without alterations in DNA sequence. Epigenetics plays a fundamental role in an array of cellular functions including development, cell proliferation and differentiation, genomic imprinting and X chromosome inactivation. Thus anomalies in epigenetic signaling could lead to tumour genesis and other health implications (reviewed in [4]). According to the classic model of cancer progression, which was proposed in 1971 by AG Knudson, two successive mutations (“hits”) are required to for a normal cell to turn into a tumour cell. In other words, two inactivation events are needed to silence the two alleles of a tumour suppressor gene [5]. However, recent studies have shown that epigenetic events may also be involved in tumour development and progression, which was initially indicated by Holliday in 1979 [6]. Epigenetic modulation of gene expression can alter tumour-related phenotype, such as invasion, angiogenesis, motility and proliferation [7]. Two most notable epigenetic alterations that take place in the mammalian genome and could impact tumour genesis are DNA methylation and histone modifications.

A.



B.

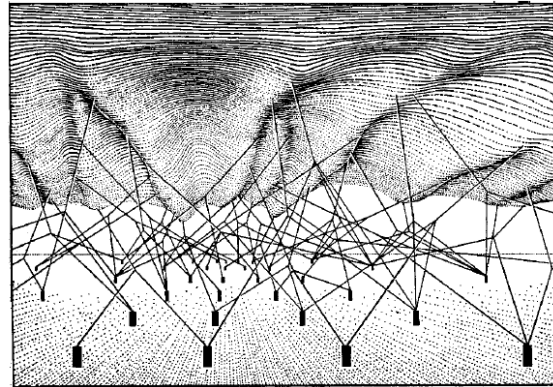


Figure 1: Waddington's illustrations

A) Waddington's epigenetic landscape. In 1957, Conrad Waddington used an epigenetic landscape to depict the decision making involved in the process of cellular development. This visual theory illustrates that the cell (represented by a ball) has specific permitted paths to follow, which lead to different outcomes or cell fates (reproduced from Waddington, p. 29)[2]. **B)** The complex system of interactions underlying the epigenetic landscape. This diagram depicts the control of the chemical tendencies of many genes on the route and the slope of any particular valley (reproduced from Waddington, p. 36) [2].

1.1.2 DNA methylation

DNA methylation is important in genome organization and transcriptional repression of unneeded genes in differentiated cells (reviewed in [7]). DNA methylation is catalyzed by DNA methyltransferases (DNMTs), which add a methyl group to the carbon 5-position of the cytosine nucleotides adjacent to a guanine (CpG), forming 5-methyl-2'-deoxycytidine (5-mdC). S-adenosyl-methionine (SAM) is the methyl donor (reviewed in [4]). In the genome, CpG di-nucleotides are usually clustered in groups or CpG islands at or near the promoters and initial exons of many genes. When these CpG di-nucleotides are unmethylated, they enable transcription. DNA methylation serves as a stable tag on the promoter of a gene and recruits other proteins such as histone deacetylases (HDACs), which form large-scale heterochromatic structures that lead to silencing of the associated genes.

1.1.3 Chromatin structure and histone modifications

The fundamental unit of a chromatin is the nucleosome, which is comprised of two pairs of core histones H2A, H2B, H3 and H4 forming a protein octomer, with approximately 146 base pairs of DNA wrapped around it (reviewed in [7, 8]) The core histones have a distinct globular fold domain that enables histone-histone and histone-DNA interaction, forming the nucleosomal disc framework. Moreover, the core histones have unstructured N-terminal sequences ("histone tails") that protrude out of the nucleosome disc and are subjected to various post translational modifications (PTMs). These PTMs include, but are not limited to, methylation, acetylation, ubiquitination, citrullination and phosphorylation (reviewed in [8]). The functional roles of these PTM

marks in transcription regulation and the mechanisms by which they modulate chromatin remodeling are now intensively studied.

1.1.4 Histone code hypothesis

Interplay between different histone PTM marks is evident and these modifications have been proposed to serve as regulatory “codes”. This is known as the histone code hypothesis, proposed in 2000 by Strahl and Allis, which states that “multiple histone modifications, acting in a combinatorial or sequential fashion on one or multiple histone tails, specify unique downstream functions” [9]. Information encoded in DNA is in part regulated by histone modifications by recruiting other proteins that either “read”, or modify the histone tails. The specific chromatin-modifying enzymes includes “writers” and “erasers”, the former adds a PTM onto the histone tail, while the latter removes the modification. Together, the “readers”, “writers”, and “erasers”, regulate the transcription of target genes [8, 10].

1.1.5 Writers, erasers and readers of the histone code

Histones are highly basic as they contain a large proportion of the two positively charged amino acids: lysine (K) and arginine (R). Therefore, modifications of these lysine residues such as acetylation or methylation cause structural alterations in the histone protein and thus the nucleosome [11]. Histone acetyltransferases (HATs) are the authors of acetylation levels; they write the acetylation mark onto these lysines, while histone deacetylases (HDACs) are the erasers that are responsible for removing the acetyl group. However, these modulations are significant only when they are relayed by histone code

readers [11]. The readers of the histone code include effector domains of proteins which bind to a specific epigenetic mark and are responsible for recruiting machinery which serve to stabilize or remodel specific chromatin states [12]. These proteins are defined by their characteristic reader module including bromodomains, chromodomains, ankyrin repeats, tudor, malignant brain tumour (MBT) domain, WD40-repeat and plant homeodomain (PHD) finger domains. Bromodomains recognize and bind to ϵ -N-acetylated lysine residues within histone tails, while the remaining modules target methylated lysines or arginines in a residue-specific manner [13, 14]. Reading of the acetylation marks could lead to further remodeling at the epigenetically modifiable sites, ultimately controlling transcriptional activity [11]. When lysines are acetylated, the positive charge on the ϵ -amino group of the side chain is neutralized, which leads to increased chromatin accessibility and thus transcriptional activation (reviewed in [15]).

Similar to acetylation, methylation too has its writers, erasers and readers. Histone lysine methyltransferases (HKMTs) and protein arginine methyltransferases (PRMTs) are the writers which catalyze the lysine and arginine methylation, respectively. Conversely, histone lysine demethylases (HKDMs) remove the methyl group from lysine residues contributing to the dynamic methylation regulation [8]. The methylated lysine and arginine residues are read by chromodomains, which were first described by Murray in 1965 [16]. Unlike histone acetylation which neutralizes the positive charge on the lysine, no such change in charge is evident with methylation. Since histone methylation does not modify the electrostatic interactions between histones and DNA, its significance for chromatin structure was undefined for decades (reviewed in [15]). However, it is now

evident that the transcriptional significance of histone methylation depends on the location of the lysine which is methylated [8].

1.1.6 Histone lysine methylation

Histones contain a high content of lysine residues which can be methylated, leading to either transcriptional activation or repression. The methylation sites that have a defined biological function include tri-methylated histone 3 lysine 4 (H3K4me3), tri-methylated histone 3 lysine 9 (H3K9me3), tri-methylated histone 3 lysine 27 (H3K27me3), tri-methylated histone 3 lysine 36 (H3K36me3), and tri-methylated histone 4 lysine 20 (H4K20me3). The H3K4me3 epigenetic mark is mostly found in promoter regions of actively transcribed genes, whereas the H3K36me3 modification is associated with transcription elongation. Moreover, the tri-methylation of H3K9, H3K27, and H4K20 are enriched at constitutively condensed chromatin and inactive genes [17]. Furthermore, in addition to being tri-methylated, lysines can be mono- and di-methylated, which augments the functional diversity of lysine methylation. As mentioned earlier, H3K4me3 is exclusive to active genes, however, H3K4me2 is found at both inactive and active genes [8].

1.1.7 H3K27 methylation

Lysine 27 on histone 3 can be mono-, di-, and tri-methylated. Tri-methylated H3K27 is an epigenetic repressive mark. H3K27me3 plays a role in development such as in X chromosome inactivation [18, 19] and parental imprinting [20] and is established by

the Polycomb group (PcG) proteins. PcG proteins exist as subunits of multi-protein complexes that are responsible for the establishment and maintenance of histone PTMs [21]. Polycomb (Pc) protein, the founding member of PcG, was initially identified as a repressor of homeotic (Hox) genes in *Drosophila*. Strong developmental defects, mostly homeotic conversions, were evident with loss of Pc function [22, 23]. Further research discovered PcG proteins to be a group of genes which, if mutated, causes phenotype reminiscent of Pc; improper body segmentation in *Drosophila*. In metazoans, at least three distinct multimeric complexes are involved in the PcG silencing pathway, which include the Polycomb Repressive Complex 1 (PRC1), PRC2, and Pleiohomeotic Repressive Complex (PHORC) [21]. PRC2 complex is extensively studied and is composed of the four core subunits. In *Drosophila*, the core subunits include Enhancer of zeste (E[z]), Suppressor of zeste 12 (Su[z]12), Extra Sex Combs (ESC) and Multicopy Suppressor of IRA1 (MSI1) [21]. These four members of the PRC2 complex are named differently in metazoans; Enhancer of Zeste Homolog 1/2 (EZH1/2), suppressor of zeste 12 (SUZ12), embryonic ectoderm development (EED) and retinoblastoma associated protein 46/48 (RbAp46/48) [24]. The enhancer of zester homolog 2 (EZH2) protein in the PRC2 complex tri-methylates H3 lysine 27 [24]. Genome-wide mapping studies have shown members of PRC1 and PRC2 complexes to co-occupy many target loci in metazoans. Furthermore, there is biochemical evidence that H3K27me3 serves as a PRC1 docking site. Together these studies are commonly interpreted as the “PcG paradigm”, which proposes that the PRC complexes act sequentially. According to this paradigm, PRC2 serves as the “initiation complex” and tri-methylates H3K27, while PRC1 complex

acts as the “maintenance complex” which recognizes the mark and subsequently mediates downstream H2A monoubiquitination [21, 25].

1.1.8 Histone lysine methyltransferases (HKMTs)

In humans, there are at least 27 histone methyltransferases, most of which are substrate specific [26]. All HKMTs, with the exception of disruptor of telomeric silencing-1-like (DOT1L) methyltransferase, have an evolutionary conserved Suppressor of Variegation 3-9 [SU(VAR)3-9], Enhancer-of-Zeste, Trihorax (SET) domain which contains the catalytic site. SUV39H1/2, G9a, GLP, and SETDB1/2 are human SUV39 proteins that use their SET domain to specifically methylate H3K9 and are related to heterochromatic gene silencing. Moreover, they contain domains that recognize epigenetic marks; SUV39H1/2 contains chromodomain, which recognizes methylated lysine and arginine residues, G9a/GLP have ankyrin repeats domain, which bind mono- and di-methylated histone 3 lysine 9 (H3K9) residues, and SETDB1/2 harbours methyl-CpG-binding domain (MBD). In addition to the SUV39 proteins, there are other HKMTs that comprise the EZH family, which are homologous to the PcG protein Enhancer of Zeste. EZH proteins methylate H3K27 with their SET domain, and unlike the SUV39 proteins, they have no Post-SET domain [8]. The Post-SET domain contains three conserved cysteine residues that play an important role in the catalytic activity of SUV39 proteins [27, 28].

HKMTs transfer a methyl group from the substrate SAM onto the terminal amine of the substrate lysine or arginine residues. With the exception of DOT1L, this transfer occurs within the SET domain, where the cofactor and peptide substrates are brought in

close proximity enabling the S_N2 transfer (nucleophilic substitution reaction) of the methyl group to occur. This reaction generates S-adenosyl-homocysteine (SAH) and the methylated histone side changes as shown in Fig. 2 (reviewed in [29]).

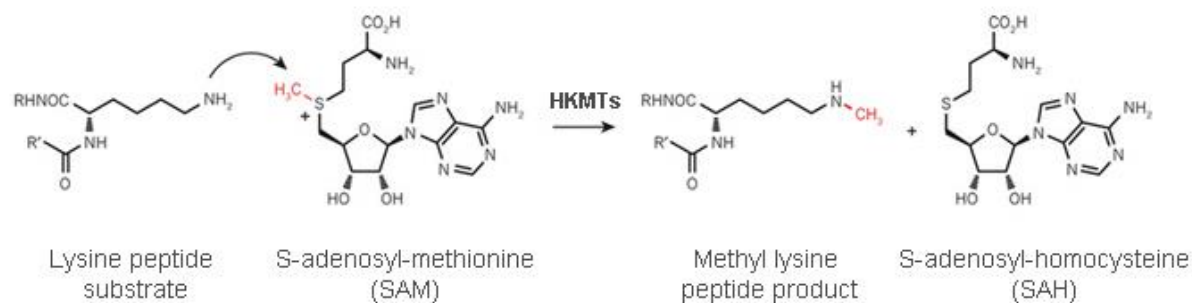


Figure 2: S-adenosyl-methionine is the methyl donor in methylation of histone residues

Histone methyltransferases transfer the methyl group from the s-adenosyl-methionine (SAM) cofactor onto the lysine amino group of the substrate histone via an S_N2 reaction. The reaction results in the formation of an N-methylated lysine and s-adenosyl-homocysteine (SAH). Figure adapted by permission from Macmillan Publishers Ltd: Nature Publishing Group, 2013 [29].

1.1.9 Recruitment of histone methyltransferases to their genomic destination

Recruitment of histone methyltransferases to specific histone targets is an important focus area and extensive research is conducted to answer the “how” and “when” of this phenomenon. Several histone-modifying enzymes are now shown to be recruited by specific DNA sequences, examples of which include the *Drosophila melanogaster* Trithorax group (TrxG) response elements (TREs) and the Polycomb group (PcG) response elements (PREs), which recruit TRX (a H3K4 methyltransferase) and PcG proteins [the polycomb repressive complex 2 (PRC2) complex containing EZH2, a H3K27 methyltransferase], respectively. TrxG is a heterogeneous collection of proteins that function antagonistically against the PcG to activate target gene expression (reviewed in [30]). Specific DNA-binding transcription factors recognize the TREs and PREs, and thus enable them to recruit these methyltransferases. A DNA sequence that enhances PcG binding is evident in human cells as well, demonstrating similarities to *D. melanogaster* PREs and suggesting that there is some conservation in the PcG recruitment mechanism.

In addition to specific DNA sequences directing methyltransferases to their histone targets, long non-coding RNAs (lncRNAs) have also been proposed to have a recruiting role. Members of the PRC2 complex, the methyltransferase G9a and the methyltransferase complex member WD repeat domain 5 (WDR5), are some of the proteins that bind to lncRNAs (reviewed in [31]). Additionally, certain lncRNAs are able to regulate multiple methyl marks. For example, the recruitment of the PRC2 complex containing the H3K27 methyltransferase and a complex containing the H3K4 demethylase KDM1A are coordinated by the human lncRNA HOTAIR to efficiently

repress specific loci (reviewed in [31]). lncRNAs may play a vast role in the specific targeting of the methyl-modifying proteins as there are 6736 lncRNAs in the human genome and can be both in cis or trans configuration which further increases the diversity of lncRNAs [32]. However, additional research is required to elucidate the mechanisms.

Small non-coding RNAs also have a role in orchestrating histone methylation and other chromatin modifications. In *Schizosaccharomyces pombe*, for instance, the establishment and maintenance of centromeric heterochromatin characterized by H3K9 methylation, requires the RNA interference (RNAi) machinery (reviewed in [31]). This is conserved in evolution to some degree, as it is also evident in *Arabidopsis thaliana* [33] and in mammals [34]. In addition to H3K9 methylation, H3K27 tri-methylation and heterochromatin changes during X-chromosome inactivation in mice are also induced by the RNAi machinery [35].

Histone methylation is also directed by DNA methylation. Some examples include the *A. thaliana* H3K9 methyltransferases SUVH4 and SUVH5 which are Set- or Ring-associated (SRA) proteins that bind to methylated DNA and also have the catalytic SET domain required for histone methylation. Mutation in the SRA domain results in significant attenuation in binding to methylated DNA *in vitro*, and reduction in DNA methylation and H3K9 di-methylation *in vivo*. This suggests that SRA domain plays a role in the recruitment and maintenance of the SUVH4 to silenced loci [36, 37].

1.2 Euchromatic human methyltransferase 2 (G9a)

1.2.1 G9a

G9a, also known as euchromatic histone methyltransferase 2 (EHMT2), was first identified in 2001 by Tachibana *et al.* [38], as a novel mammalian lysine-preferring HMTase. G9a methylates the lysine residues of histone H3 like SUV39, the first identified lysine-preferring HMTase, but with a 10-20 fold increased activity. In addition, compared with the SUV39 proteins which methylate only H3K9, G9a methylates H3K9 and to a lesser degree, H3K27 [38]. Moreover, while SUV39 proteins tri-methylate H3K9 specifically at the pericentric heterchromatic region, G9a is responsible for mono- and di-methylation of lysine 9 on the euchromatic loci involved with transcriptionally active genes [39]. Also, G9a methylates histone H1.4, and non-histone proteins, including p53 (K373) [41], WIZ (K305) [42], CDYL1 (K135) [40, 42], ACINUS (K654) [42], Reptin (K67) [43], and itself at K165 [40]. G9a auto-methylation plays a role in G9a-HP1 interaction [40, 44]. The biological significance of the methylation of these non-histone proteins remains to be elucidated.

1.2.3 G9a structure

G9a exists as a heteromeric complex with the highly similar methyltransferase G9a-like protein [GLP, also known as euchromatic histone methyltransferase 1 (EHMT1)]. Mono- and di-methylation of H3K9 is greatly attenuated when either G9a or GLP methyltransferase activity is diminished [45]. Double knockout of G9a and GLP

does not augment the reduction in H3K9me1 and H3K9me2 levels [46]. Thus, the loss of GLP activity *in vivo* generally cannot be compensated by G9a, and vice versa.

G9a consists of an N-terminal auto-methylation site, a glutamic acid-rich (E-rich) domain, a cysteine-rich (Cys-rich) domain, ankyrin repeats and a SET domain [47] (Fig. 3). The N-terminal of G9a contains an auto-methylation site, at lysine 165, which can be di-methylated [40]. Methylated G9a binds and recruits histone protein 1 (HP1) [40, 44]. The E-rich region of G9a composed of 24 contiguous glutamic acid residues (300-323aa) [47] is similar to the poly-acidic regions present in nuclear protein nucleolin [48], the non-histone chromosomal high mobility group 1 (HMG1) protein [49] and the major centromere auto-antigen CENP-B [50, 51]. The Cys-rich region of G9a contains 12 cysteine residues. The Cys-rich region acts as a binding site for neuron-restrictive silencing factor (NRSF), via which G9a is involved in repression of neuronal genes in non-neuronal tissues [52].

The ankyrin repeats of G9a (649-879aa) are shown to be involved in epigenetic mark recognition. They bind to mono- and di-methylated lysines and their specificity is comparable to that of other methyl binding protein modules including chromodomains, tudor domains and PHD fingers. Ankyrin repeats are helix-turn-helix- β -turn structures with the helices stacked and the β -turns projected out at right angles. The β -turns and the helices of the fourth and fifth repeats sandwich the H3 peptide, while a partial hydrophobic cage formed by three tryptophans (Trp839, Trp844 and Trp877) and one acidic residue (Glu847) recognize and bind the di-methylated lysine (H3K9me2) epigenetic tag. The methyl groups of H3K9me2 project towards the hydrophobic

tryptophans while the proton of the N ϵ atom of the lysine bridges to the acidic Glu847, and thus mutations in these G9a residues attenuate peptide binding. This hydrophobic pocket binds H3K9me1 and me2 with equal affinity; however, it does not bind H3K9me3 due to steric hindrance [53].

The SET domain is conserved among histone methyltransferases and contains the catalytic methyltransferase activity. The SET domain (1038-1155aa) has a pseudo-knot-type structure comprised of a β -sheet fold separated by an α -helical iSET motif [54, 55]. It contains a central channel which binds SAM and the protein substrate on opposite sides of the domain. The methyl group of SAM is positioned at the base of the channel, while the side chain of the target lysine protrudes into the channel from the other end. The substrates within the channel are held in position via hydrogen bonding which allows the methyl group from SAM to be transferred onto the lysine ϵ -amine by a S_N2 reaction [56, 57].

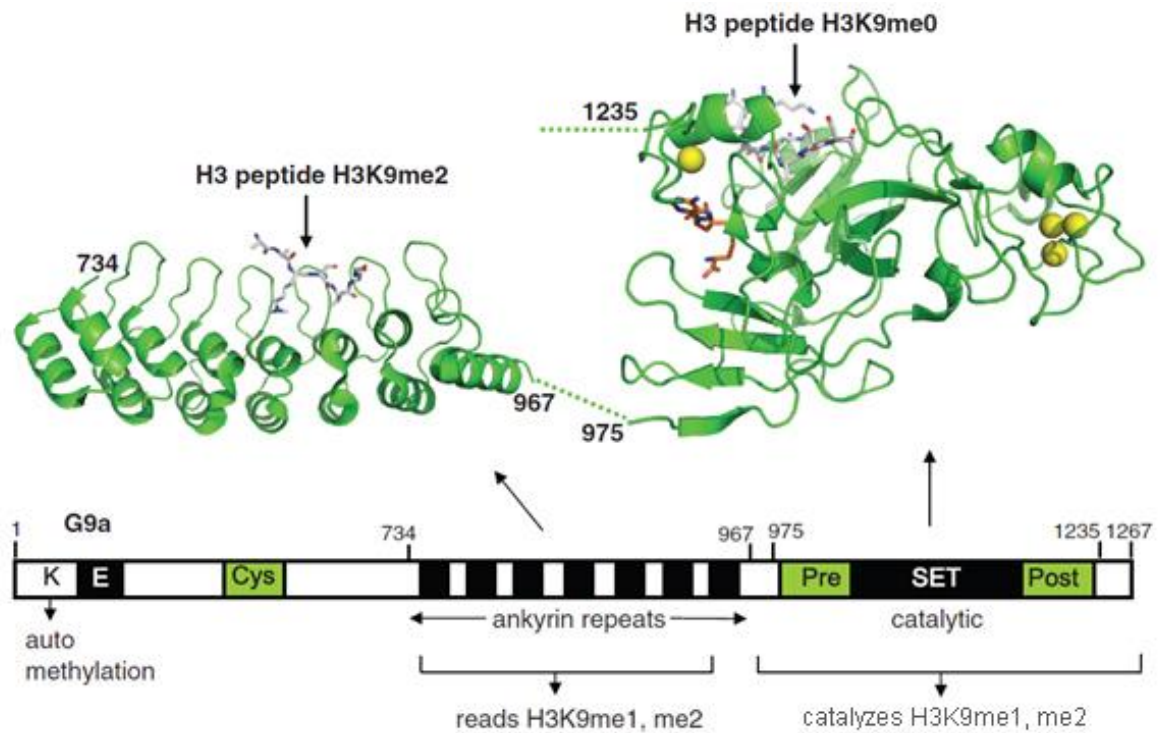


Figure 3: Domain structure of G9a

G9a consist of an N-terminal automethylation site (K165), E-rich domain made of 24 contiguous glutamic acid residues, Cys-rich domain comprised of 12 cysteine residues, the product-binding ankyrin repeats and the catalytic SET domain. Adapted by permission from Oxford University Press: Nucleic Acids Research, 2010 [45].

1.2.4 The effect of other post-translational modifications on G9a activity

G9a methylation is not dependent on pre-existing epigenetic marks like some other HKMTs including SET1 or DOT1 [58]. The absence of the epigenetic marks on the target lysine or adjacent residues is sufficient for G9a to transfer the methyl group. G9a interacts with H3 at amino acids 6-11, with a seven amino acid heptapeptide RK/ARK consensus, such as 6-TARKSTG-12 of H3, being the minimum substrate recognition criterion *in vitro* [59]. When H3K9 is acetylated, or when H3S10 or H3T11 are phosphorylated, H3K9 methylation is blocked. H3K4 acetylation or methylation, on the other hand, reduces H3K9 methylation. Moreover, H3R8 methylation blocks H3K9 methylation [42, 59]. Phosphorylation of H3S10 also blocks product binding [53].

1.2.5 Recruiting other proteins to reinforce silencing

G9a acts as a chromatin-compacting gatekeeper and this role is further reinforced by its ability to recruit other proteins to chromatin. These proteins include histone H1.4 [60], heterochromatin protein 1 (HP1) [40, 44] and CCAAT/ Enhancer-binding protein- β (C/EBP- β) [61]. Histone H1 has 11 variants in human cells, all of which are generally silencing and play a role in establishing higher order chromatin structure. Nucleosome repositioning and sliding is blocked by H1 arrangement. H1 positioning also prevents DNA accessibility to other proteins (reviewed in [45, 62]). H1.4 of the H1 variants contributes to dynamic, gene specific silencing [60]. G9a methylates H1.4 at lysine 26 (H1.4K26) and plays a direct role in recruiting H1.4 to chromatin [60]. H1 methylation further reinforces silencing as it acts as another epigenetic mark to recruit HP1 [60, 63],

L3MBTL1, which participates in chromatin compaction [60, 64] and MSX1, a homeobox transcriptional repressor [65].

G9a and GLP are essential in recruiting HP1 to euchromatin. When the activity of either G9a or GLP is attenuated, HP1 recruitment to euchromatin binding is diminished; HP1 occupancy at specific promoters is reduced with loss of G9a [46, 66, 67]. G9a recruits HP1 by auto-methylating, possibly at multiple sites [40, 44]. The dominant auto-methylation site (K165) of G9a resembles the region surrounding H3K9. There is even a phosphorylation site that diminishes HP1 binding. When the target lysine (K165) is mutated, the HP1 binding is eliminated and it delocalizes from G9a-containing loci. In addition to binding to the auto-methylated G9a methyltransferase, HP1 binds specifically to di-methylated and tri-methylated H3K9 by its N-terminal chromodomain (reviewed in [68]). HP1 recruits other proteins and DNA methyltransferases (DNMTs) not only in the formation of silent heterochromatin but also euchromatin (reviewed in [69]). A direct interaction between G9a and the DNA methyltransferases has been proposed by some studies [70, 71]. Nonetheless, HP1 binding to auto-methylated G9a or H3K9me₂/me₃ could play a role in recruiting the DNA methyltransferases DNMT3a and DNMT3b. Certainly, DNMT3a, DNMT3b and CpG methylation is attenuated in the absence of HP1 or G9a, while HP1 knockdown has little effect on G9a binding at the TNF α promoter. This indicates that G9a is the main player of silencing at that promoter, while HP1 acts as a bridge serving to recruit DNMTs [67].

In addition, studies have shown G9a to bind and methylate CCAAT/Enhancer-binding protein- β (C/EBP- β). C/EBP- β methylation blocks its ability to activate

transcription [61]. When the catalytic SET domain of G9a is not present, the binding is not observed. This suggests product binding is the active mechanism, but this possibility has not been formally tested [61]. Overall, G9a interacts with other proteins and recruits them to target genes to reinforce silencing.

1.2.6 The role of G9a: biological relevance

Multiple biological roles for G9a have been proposed since it is ubiquitously expressed and functions as the major euchromatic H3K9me1 and H3K9me2 histone methyltransferase (HKMT). G9a is essential in mouse development; G9a knockout results in embryonic lethality around embryonic day 9.5 (E9.5) due to severe growth defects [46, 72]. This is further supported by the fact that mutant mice with catalytically inactive G9a HKMT also have embryonic lethal phenotype, as seen in the G9a knockout mice (reviewed in [73]). When G9a is knocked out in cells, DNA methylation and HP1 binding to euchromatin is diminished, which results in loss of silencing [72]. In addition, differentiation of ES cells in absence of G9a leads to apoptosis [72]. Together, these data suggest that H3K9 methylation by G9a is important for mouse development.

Furthermore, developmental defects of germ cells and lack of meiosis completion in either sex is seen in germ lineage-specific G9a knockout [74]. When G9a is specifically knocked out in lymphocytes, lymphocyte development is scarcely affected, with some impairment in the immune responses. A reduced usage of Ig λ L chains and a diminished Ig λ gene assembly in bone marrow precursor cells is evident in B-cell-specific G9a knockout mice. B cells deficient in G9a do not respond well to LPS + IL-4 *in vitro* [75]. T-cell-specific G9a knockout mice have normal T-cell development in the

thymus; however there is impaired expression of lineage-specific Th2-associated cytokine genes such as IL-4, IL-5, and IL-3. Thus, the mice are unable to defend themselves against parasitic infections [76].

Neuron-specific G9a knockout mice have interesting postnatal and forebrain phenotypes [77]. These mice have multiple non-adult neuronal genes de-repressed in the forebrain but with no apparent neuronal developmental or architectural defects. Nonetheless, various abnormal phenotypes, especially defects in cognition and adaptive behaviour, are evident in these mice. G9a has a similar role in neuronal function in *Drosophila*, where G9a deficiency results in learning and memory defects [78]. Furthermore, G9a-mediated H3K9me2 methylation has been reported to play a critical role in cocaine-induced neuronal responses. Alterations in gene expression in the nucleus accumbens (NAc) neurons in mice is seen with repeated cocaine exposure. These alterations are critical for cocaine-induced changes in neural morphology and behaviour that may be responsible for cocaine addiction. One of the changes that occur with repeated cocaine exposure in mice is the reduced expression levels of G9a and GLP specifically in the NAc neurons, leading to reduced H3K9me2 levels. This attenuation in H3K9me2 levels subsequently causes up-regulation of G9a-repressed genes. G9a (G9a-GLP complex) or the methyltransferase activity of G9a may be a novel target for regulating cocaine addiction if these molecular mechanisms in mice can be generalised to cocaine addiction in humans [79].

Additionally, G9a is required for the silencing of the homeobox gene Oct3/4, which in turn is important for pluripotency [66]. It is also important in NRSF/REST-

mediated silencing of neuronal genes in non-neuronal lineages [52], PRDI-BF1 silencing in B-cell differentiation [80] and CDP/cut mediated silencing of genes essential in cell differentiation and proliferation [81].

Given the role of G9a in various processes, including cell differentiation and proliferation [81], it is not surprising that vicissitudes in G9a expression level are seen in many cancers. The expression of G9a is up-regulated in many tumors including hepatocellular, colon, prostate, lung and invasive transitional cell carcinomas and in B cell chronic lymphocytic leukemia [41]. In bladder and lung cancer cells, growth suppression and apoptosis with an increase in sub-G1 population are seen with G9a knockdown [82]. High levels of G9a correlate with poor prognosis in aggressive lung cancer, with augmented *in vitro* cell migration and invasion, and *in vivo* metastasis. The increased metastasis that is evident in lung cancer is due to G9a repression of the epithelial cell adhesion molecule, Ep-CAM [83]. In addition, G9a promotes EMT by repression of E-cadherin expression in Claudin-low breast cancer (CLBC), which is associated with early metastasis [84]. G9a knockdown de-repressed E-cadherin expression, preventing cell migration and invasion *in vitro*, and attenuated tumour growth and metastasis in CLBC metastasis models [84].

1.2.7 G9a inhibitors

Numerous inhibitors have been developed through structure-based design, targeting the catalytic domains of lysine methyltransferases. The two potential sites for inhibition include the SAM cofactor binding site and the protein substrate binding cleft. Sinefugin, chaetocin, and SAH are some of the natural products that act as competitive

inhibitors of SAM. Sinefungin and SAH are not KMT-specific and are structurally similar to SAM, while chaetocin's specificity remains questionable. Chaetocin was first described to be a selective inhibitor of *Drosophila* SU(VAR)3-9, a homologue of human SUV39H1; however, recent studies have shown this specificity to be diminished in certain *in vitro* methylation assays (reviewed in [57]). The second class of inhibitors that act on the protein substrate binding cleft of lysine methyltransferases include BIX-01294 and UNC0224. BIX-01294 was among the first of this type of inhibitor to be designed and it is a specific G9a and GLP inhibitor (reviewed in [57]). BIX-01294 reduces the levels of H3K9me2 in cells and also promotes stem-cell reprogramming with similar efficacy as the viral transduction of pluripotency-inducing transcription factors [85]. BIX-01294 binds the substrate binding cleft in similar manner as observed in the histone H3 peptide complex, forming similar hydrogen bonds and van der Waals contacts. The quinazoline ring of BIX-01294 binds in a region analogous to that bound by the backbone atoms of residues T6-R8 and the side chain of R8 in histone H3 peptide. The R8 residue is critical for G9a activity and BIX-01294's similar conformation as the R8 residue enables it to efficiently inhibit the methyltransferase [42]. Moreover, the diazepane ring of BIX-01294 also binds to the same location as the side chain of K4, and the benzene and piperidine groups occupy similar positions of residues Q5 and T6, respectively, further enhancing the binding affinity of the inhibitor [42].

Derivatives of the quinazoline scaffold of BIX-01294 were made by Liu *et al.* to develop the more potent G9a inhibitor UNC0224 with a K_d value of 23 nM, which has approximately fivefold greater affinity than that of BIX-01294, which has a K_d value of

130 nM [86]. Recently, further optimizations to UNC0224 have been made, yielding even more efficient inhibitors with at least sevenfold potency compared to BIX-01294 [87]. These inhibitors are useful in studying the role of G9a and also illustrate the possibility of such structure-based inhibitor design for other oncogenic SET-domain lysine methyltransferases such as EZH2 and SMYD3 [57].

1.3 Chromodomain, Y-like: CDYL

1.31 CDYL

A group of very closely related genes on the human Y chromosome named *ChromoDomain Y(CDY)*, was discovered in 1997 by Lahn *et al.* [88]. In 1999, Lahn *et al.* [89] identified an autosomal homolog of *CDY*, *CDYL (CDY Like)* in humans, and *Cdyl* in mice. Human *CDYL* and *CDY* are paralogs, and their protein products share 63 % identity. On the other hand, human *CDYL* and mouse *Cdyl* are orthologs, having protein products that share 93 % identity. Using comparative evolutionary studies, Lahn *et al.* (1999), showed that *CDY* genes are a result of retroposition of *CDYL* mRNA and amplification of the retroposed gene on the Y. Mice do not have *CDY* on the Y chromosome. Instead, mice have two autosomal *Cdyl* gene transcripts, a highly abundant testis specific 2.8 Kb transcript and a ubiquitous 3.6 Kb transcript [89]. The mouse *Cdyl* has been implicated in spermatogenesis. It is abundantly expressed in elongating spermatids during histone hyperacetylation and replacement. In humans, *CDY* gene encodes the chromodomain Y protein, which is specifically expressed in the testes, whereas *CDYL* expresses the ubiquitous CDYL transcript [89].

1.3.2 CDYL domains and function

Both *CDY* and *CDYL* contain a canonical chromodomain in their N-terminal region and domain with significant sequence homology to the enoyl-coenzyme A hydratase/isomerase catalytic domain near the C-terminal region [89]. The mouse *Cdyl* recruits HDAC1 and HDAC2 and binds to coenzyme A (CoA) via its putative catalytic

domain [90]. However, HDAC recruitment represses the catalytic domain's ability to bind CoA. Thus, Cdy1 interacts with CoA only in the absence of HDACs. Caron *et al.* (2003) [90] found that the cellular content of HDAC1 and HDAC2 decreases significantly in mouse spermiogenesis, while Cdy1 levels do not. This decrease in HDAC1 and HDAC2 levels would enable Cdy1 to bind CoA. In 2002, Lahn *et al.* [91] claimed recombinant Cdy1 to have *in vitro* HAT activity, and they suggested Cdy1 to play a role in histone acetylation during spermiogenesis. However, Caron *et al.* (2003) [90] were not able to detect any significant HAT activity associated with recombinant Cdy1. Thus, the role of Cdy1 in histone acetylation during elongation of spermatids remains unclear.

CDYL has been identified to be a component of repressor complexes C-terminal binding protein (CtBP) [92] and REST/CDYL/G9a [93], where CDYL bridges G9a to the transcription repressor RE1-binding silencer of transcription (REST). Mulligan *et al.* (2008) [93] showed CDYL to also interact with mesoderm early response 1 (MIER1), HDAC1 and HDAC2. Since CDYL is part of such repressor complexes, CDYL has been suggested to mainly function as a transcription corepressor in somatic cells [90, 94, 95]. CDYL contains a chromodomain near its N-terminus, which generally recognizes methylated histone lysine residues [89]. CDYL specifically interacts with the repressive H3 lysine methylation marks, H3K9me3, H3K27me2 and H3K27me3, via its chromodomain. CDYL does not bind H3K4 methylation modifications which are associated with transcriptional activation [96]. Furthermore, CDYL interacts with PRC2 complex proteins, SUZ12, EED and EZH2 *in vivo* [96]. CDYL directly interacts with EZH2 and increases the activity of EZH2 towards oligonucleosome substrates. Genome-

wide analysis showed CDYL to share a significant number of common genomic targets with EZH2. However, CDYL does not bind to classical PRC2 target promoters such as those of the HoxA family genes [96]. HoxA is one of the four clusters of Hox family genes (reviewed in [97]). Even though Zhang *et al.* (2011) [96] showed CDYL to interact with EZH2, previous biochemical purification approaches did not show CDYL to be in a complex with the PRC2 complex [98]. Thus, it is possible that CDYL is a peripheral component of the PRC2 complex. Further research is required to determine the precise role of CDYL in the biological functions of PRC2 complex and H3K27 methylation.

1.4 Mesoderm induction early response 1: MIER1

1.4.1 MIER1

Mesoderm induction early response 1 (MIER1) is a fibroblast growth factor (FGF)-activated transcriptional regulator originally identified in *Xenopus laevis* [99]. It is highly conserved in evolution, with 95 % identity between human and mouse sequences at the amino acid level. The *mier1* gene spans 63 Kb and consists of 17 exons. It produces six protein isoforms with different amino and carboxyl termini [100], which have a very low or undetectable expression in healthy adult tissues [100, 101]. However, there is high expression of these isoforms in human breast carcinoma cell lines and breast tumours [100]. Moreover, MIER1 α expression pattern analysis using breast samples from patients showed a dramatic shift in the subcellular localization in ductal epithelial cells. A shift from nuclear to cytoplasmic localization was evident during progression to invasive carcinoma [102]. Thus, regulating nuclear levels of MIER1 α may play an important role in the development of invasive breast cancer.

1.4.2 MIER1 structure and function

The *mier1* gene encodes multiple distinct MIER1 proteins through exon skipping, facultative intron usage, alternative polyadenylation sites and the use of two alternate promoters, P1 and P2 [101]. The MIER1 proteins have common internal region, but distinct amino (N-) and carboxy (C-) terminal sequences. Thus far in humans, MIER1 isoforms with three different N- and two different C- termini have been characterized. The N-terminal isoforms result from the use of the two alternate promoters and

alternative splicing [101]. Even though P1 and P2 transcription results in mRNAs with distinct 5'UTRs, with the exception of the first two amino acids after the start methionine, the resulting protein sequences (named N2 and N3 in GenBank) are identical. Transcription only from the P1 promoter allows for alternative inclusion of a cassette exon, exon 3A. When this exon is included, it transcribes the third N-terminal isoform (named N1 in GenBank) with additional sequence at its N-terminus containing a *bona fide* nuclear export signal (NES) [103].

The two C-termini isoforms, α and β , are produced by the alternate use of a facultative intron at the 3'end and differ significantly in size and sequence. As demonstrated by PCR analysis, the β isoform is more widely expressed and abundant than the α isoform [101]. MIER1 α C terminus results from the removal of the facultative intron during RNA processing and contains 23 amino acids (aa). The α C terminus includes the LXXLL motif, which is essential for interaction with nuclear hormone receptors [101, 104]. MIER1 α interacts with ER α in MCF7 cells and MIER1 α overexpression inhibits estrogen-stimulated anchorage-independent growth [102]. The MIER1 β C terminus results from inclusion of the facultative intron and consists of 102 aa, including the functional nuclear localization signal (NLS) [101]. The two isoforms' distinct C termini amino acid sequences suggest different functions.

MIER1 isoforms have a number of structurally similar features as do other transcriptional regulators. These include the acid activation domain, an ELM2 domain, and a SANT domain, as shown in Fig. 4 [99]. The ELM2 (EGL-27 and MTA1 homology)

domain was initially defined in Egl-27, a *Caenorhabditis elegans* protein that has an important role in patterning during embryonic development [105]. The ELM2 domain is evolutionarily conserved. In MIER1, the ELM2 domain serves to recruit histone deacetylase 1 (HDAC1) and repress transcription [106]. In addition, MIER1 binds to CREB-binding protein (CBP), inhibiting its histone acetyltransferase (HAT) activity. The N-terminal region of MIER1, including both the acidic and ELM2 domains, is responsible for this interaction [107].

The SANT domain was first identified in the transcription factors SWI3, ADA2, N-CoR and TFIIB, which gave the acronym “SANT” [108]. The SANT domain is found in other proteins as well, including the transcriptional regulatory molecules involved in nuclear hormone activity, N-CoR [108] and SMRT [109], the components of transcriptional chromatin-regulatory complexes, MTA-1 [110] and MTA-2 [111] and the proteins essential for regulating developmental events, Egl-27 [105] and CoREST [112]. In addition to being implicated in DNA binding, the SANT domain has been shown to be involved in protein-protein interactions [108]. It has been shown to interact with complexes containing histone deacetylase (HDAC) [112, 113] and histone acetyltransferase (HAT) [114]. Nonetheless, the precise role of the SANT domain in individual proteins is unknown. In MIER1, the SANT domain interacts with Sp1, and dissociates it from its binding sites in the promoter of regulated genes, including MIER1’s own promoter and interfering with transcription [115]. There is an implied structural and/or functional relationship between the SANT and the ELM2 domain since most ELM2-domain containing proteins also have a SANT motif [103]. Furthermore, MIER1

interacts with G9a, a histone methyltransferase [116], but the effect MIER1 has on this chromatin modifying enzyme is unknown. Overall, both isoforms of MIER1 can function as transcriptional regulators by their interaction with chromatin modifying enzymes.

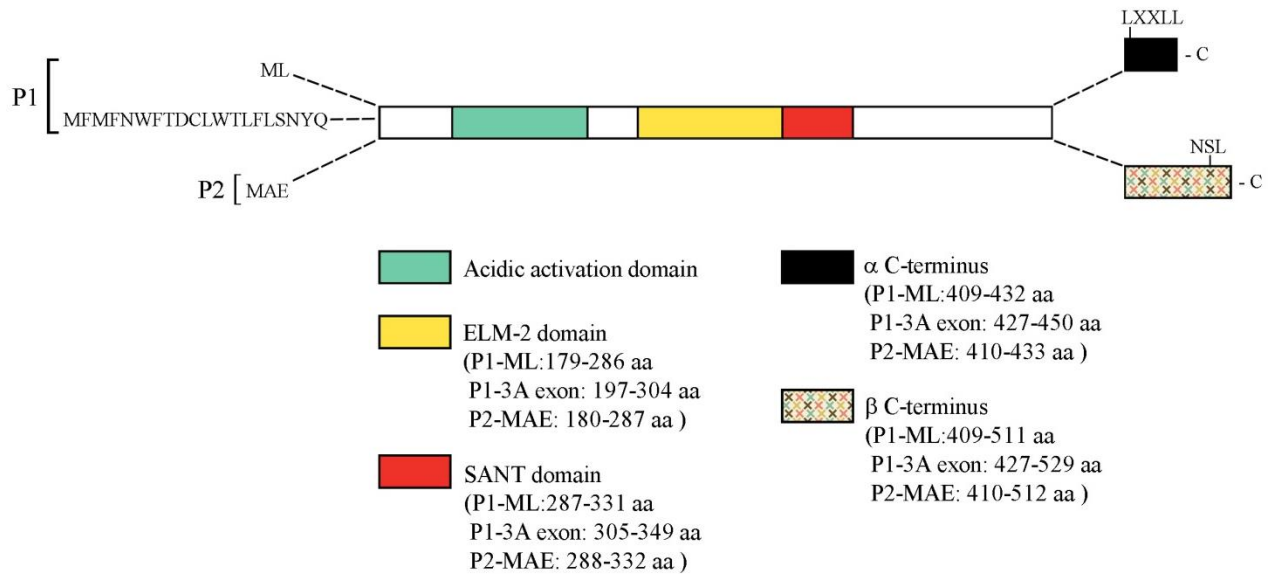


Figure 4: Domain structure of MIER1

MIER1 has two major isoforms, MIER1 α and MIER1 β . These two C-termini isoforms are produced by the alternate use of a facultative intron 16 at the 3' end; MIER1 β C terminus results from the inclusion of this facultative intron. MIER1 α C-terminus contains a classic LXXLL motif for interaction with nuclear receptors, and MIER1 β C-terminus contains a nuclear localization signal (NLS). Two alternative promoters, P1 and P2, result in distinct MIER1 transcripts as well. Transcription from the P1 promoter produces MIER1 isoforms that either begin with M-L- or MFMFNWFTDCLWTLFLSNYQ- (sequence encoded by exon 3A). Transcription from the P2 promoter produces MIER1 isoforms that begin with MAE-. All MIER1 isoforms share a common set of domains; the acidic domain, SANT domain, and the ELM-2 domain, as shown above.

1.5 Hypothesis and objectives

MIER1 has been shown to interact with chromatin modifying enzymes and to potentially act as a transcriptional regulator. There has also been recent evidence illustrating the role of MIER1 in specifically recognizing and binding to acetylated H3K27 (Paterno, unpublished data). Furthermore, MIER1 interacts with HDAC and recent experiments have shown MIER1 to augment HDAC activity towards acetylated H3K27, removing the acetyl group from lysine 27 (Paterno, unpublished data). Moreover, MIER1 has been shown to interact with G9a, a histone methyltransferase [116]. However, the MIER1-G9a interaction has been only weakly examined. **Taking these together, the purpose of this study was to investigate the interaction between G9a and MIER1 and to determine their role in H3K27 methylation. We hypothesize MIER1 interacts with G9a directly, and we also hypothesize MIER1 complexes with a methyltransferase to methylate H3K27 (Fig. 5).**

Objective 1: Determination of a possible direct interaction between MIER1 β and G9a

GST pull-down assay will be performed between GST-MIER1 β and various GST-MIER1 β deletion constructs, and recombinant G9a to determine the possibility of an interaction between G9a and MIER1 β *in vitro*.

Objective 2: Characterization of the *in vivo* interaction between MIER1 α and G9a

Human embryonic kidney cells (HEK-293) will be transfected with MIER1 α and various constructs of G9a, and coimmunoprecipitations will be conducted in order to determine whether MIER1 interacts with G9a *in vivo*.

Objective 3: Determination of possible interaction between MIER1 α , G9a and CDYL

Human embryonic kidney cells (HEK-293) will be transfected with MIER1 α , G9a, and CDYL, and co-immunoprecipitations will be conducted in order to determine whether they exist as a complex *in vivo*.

Objective 4: Determination of possible interaction between MIER1 α and EZH2 in the presence or absence of CDYL

Human embryonic kidney cells (HEK-293) will be transfected with MIER1 α and CDYL, and co-immunoprecipitation will be conducted in order to determine whether MIER1 α interacts with EZH2 in the presence of CDYL.

Objective 5: Determination of possible H3K27 methylation by recombinant G9a

The activity of recombinant SET domain and full length G9a towards H3K27 methylation will be tested with and without the addition of inhibitors (SAM and BIX-01294), and detected using the streptavidin coated plates and ELISA approach, along with Western Blotting.

Objective 6: Determination of GST-MIER1 β methylation activity on H3K27.

The activity of GST protein and GST fused MIER1 β protein towards H3K27 methylation will be tested with and without the addition of inhibitors (SAM and BIX-01294), and detected using the streptavidin coated plates and ELISA approach, along with Western Blotting.

Objective 7: Determination of H3K27 methylation by MIER1 α complex

Human embryonic kidney cells (HEK-293) will be transfected with myc-MIER1 α , and immunoprecipitation will be conducted in order to pull down transiently

expressed protein, along with its *in vivo* binding partners. The proteins will then be used in the methylation assay, and the methylation activity will be detected using streptavidin coated plates and ELISA approach.

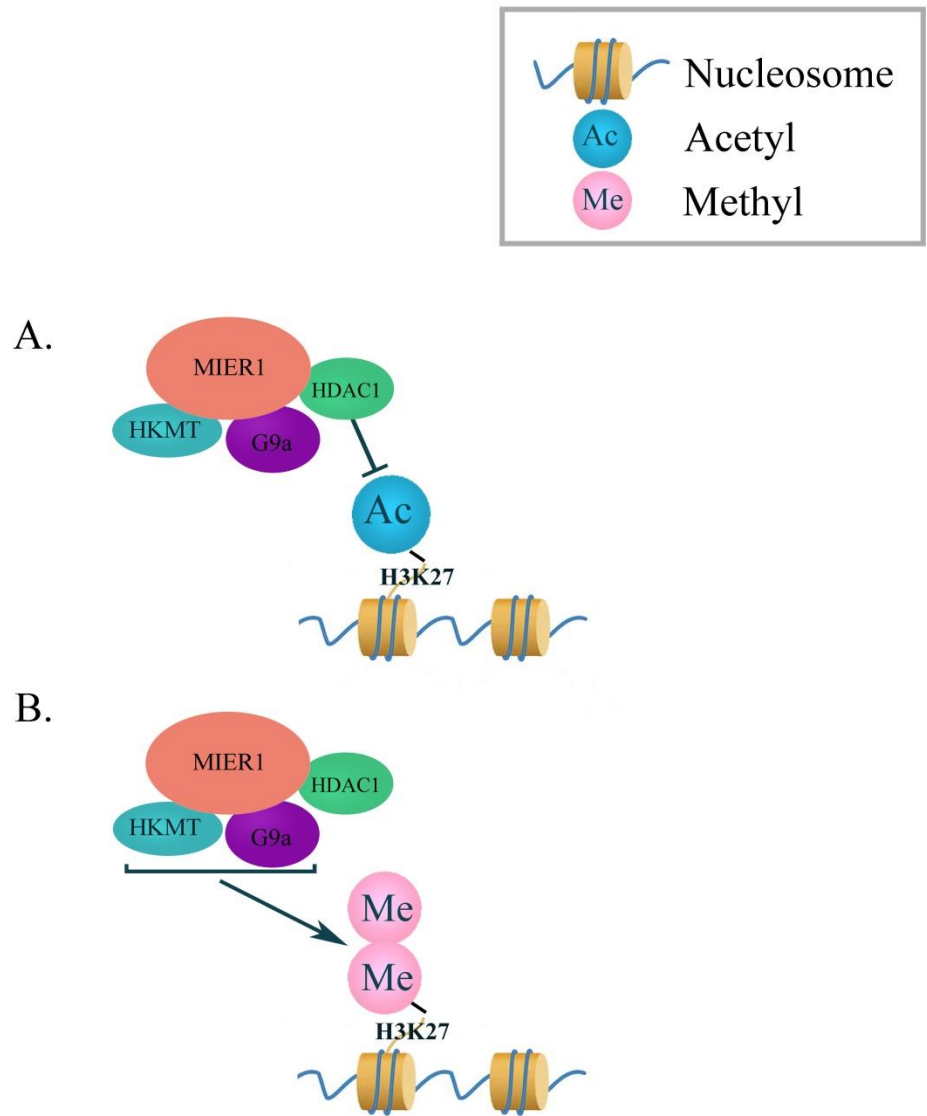


Figure 5: Hypothesised role of MIER1 in H3K27 methylation

MIER1 acts as a transcriptional regulator by interacting with chromatin modifying enzymes such as HKMTs and HDACs. Recent evidence has shown MIER1 to specifically bind to acetylated H3K27 (Paterno, unpublished data). **A.** MIER1 augments HDAC1 activity towards acetylated H3K27, removing the acetyl group from lysine 27 (Paterno, unpublished data) and subsequently enabling H3K27 to be methylated. **B.** MIER1 complexes with G9a and/or another putative HKMT and methylates H3K27.

Chapter 2: Materials and Methods

2.1 Plasmids

A. GST-MIER1 β , GST-aa1-283, GST-aa287-512, GST-aa287-357

MIER1 β (GenBank: AAM97503.1) and its deletion constructs were cloned by Zhihu Ding into the pGEX-4T-1 plasmid (a kind gift from Dr. A. Pater, Memorial University), which contains the GST-tag, at the EcoRI restriction site [115].

B. GST-SANT mutant

GST-MIER1 β with three point mutations at amino acids ³²⁶FYY→AAA was constructed by Julia Ferguson using QuikChange II XL site-directed mutagenesis kit (Stratagene 200521).

C. pCS3+MT, myc-MIER1 α

pCS3+MT is a myc-tagged empty vector, which was a gift from Drs. David Turner and Ralph Rupp (University of Michigan). Corinne Mercer cloned MIER1 α (GenBank: AAM97500.1) into this vector at the BglII restriction site, making myc-tagged MIER1 α [101, 106].

D. myc-G9a-SET

Myc-G9a-SET was cloned into pCS3+MT vector between EcoR1 and NotI restriction sites from EHMT2/pET28a-LIC plasmid (Addgene 25503), via PCR cloning. EHMT2/pET28a-LIC plasmid was used to clone the SET domain into pCS3+MT vector. The EHMT2/pET28a-LIC plasmid purchased from Addgene was a bacterial expression with resistance to kanamycin. It was plated on kanamycin plate and incubated overnight at 37 °C. A bacterial colony was picked and grown overnight at 37 °C in 250 ml of lysogeny broth (LB) with kanamycin (50µg/ml) on a shaker. The plasmid was prepared using the Endotoxin-Free Maxi Plasmid Kit (Qiagen 12362).

PCR amplification of SET domain from the pET28a-LIC plasmid

PCR was performed on the EHMT2/pET28a-LIC plasmid using G9a-SET primers, G9a-SET Forward primer (5' GCACCATGGGCAGCCATCATC 3') and G9a-SET Reverse primer (5' GTGAATTCTCAAAACCCTAGCTCCTCCC 3'), as listed in Table 1. PCR reagents and setting used are shown in Tables 2 and 3, respectively. The PCR reagents were mixed and the PCR was performed to amplify the SET domain. The PCR product (2 µl) was run on a 1.2 % agarose gel with SYBR Safe DNA Gel Stain (1 in 10,000 dilution) (10,000X, Invitrogen S33102) to verify the presence of the amplified SET domain. After amplification with Platinum *Taq* DNA polymerase High Fidelity, the addition of 3'A-overhangs was performed. The vial was placed on ice and 1 unit of *Taq* DNA polymerase (0.2 µl) was added to the PCR product. It was mixed and incubated at 72 °C for 10 min. It was then placed on ice and used immediately in the TOPO TA cloning reaction as follows using the TOPO TA Cloning Kit (Invitrogen 45-0641): the

reagents – 4 µl of PCR product, 1 µl of salt solution (1.2 M NaCl; 0.06 M MgCl₂), 1 µl of TOPO vector and 5 µl of dH₂O – were gently mixed and incubated for 5 min at room temperature and then placed on ice.

Following the TOPO TA cloning reaction, One Shot Mach1-T1R competent cells (Invitrogen 45-0641) were transformed. TOPO TA cloning reaction product (2 µl) was added into a 1.7 ml sterile tube containing One Shot chemically competent cells and mixed gently. It was then incubated on ice for 30 min. The cells were then heat shocked for 30 sec at 42 °C and were then immediately transferred on ice. Room temperature LB medium (250 µl) was added to the cells and incubated at 37 °C for 1 hr on a shaker. The transformation (100 µl) was spread on a pre-warmed agarose plate with ampicillin selection and incubated overnight at 37 °C. Thirteen colonies were selected the following day and a PCR was performed using “G9a-SET Forward” and “G9a-SET Reverse” primers listed in table 1 to check for the presence of the SET domain in the TOPO vector. The reagents and setting used for PCR analysis are shown in Tables 4 and 5, respectively. The PCR products were run on 1.2 % agarose gel with SYBR Safe DNA Gel Stain (1 in 10,000 dilution) (10,000X, Invitrogen S33102) to visualise the products. Colony containing the SET domain/Topo vector was grown in 6 ml of LB with ampicillin (Amp) overnight at 37 °C. Miniprep was performed on the culture using the Promega Miniprep Kit (A1460), following the manufacturer’s protocol. Nanodrop Spectrometry (Thermo Scientific ND1000) was used following manufacturer’s protocol, to determine the concentration of the plasmid yielded.

Table 1: PCR primers

Gene of Interest	Nucleotide sequence	Size of Amplicon
Full length G9a Forward	5'CGGAATTCTATGGGGGCGCTCCTGCTGGAGAA G 3'	3,570 bp
Full length G9a Reverse	5'CGCTCGAGCCGTTCTCATGTGTTGACAGGGGG 3'	
G9a-N-terminal Forward	5'CGGAATTCTATGGGGGCGCTCCTGCTGGAGAA G 3'	1,824 bp
G9a-N-terminal Reverse	5'CGCTCGAGTCACGGAGCTTCTTCCGCCTCTCTG AC 3'	
G9a-ANK Forward	5'CGGAATTCTATGTTTCCACCCTCGGCAGTTGTA CCT 3'	693 bp
G9a-ANK Reverse	5' CGCTCGAGTCACAGCTCAGGGTTGGCCCC 3'	
G9a-SET Forward	5' GCACCATGGGCAGCCATCATC 3'	690 bp
G9a-SET Reverse	5' GTGAATTCTCAAAACCCTAGCTCCTCCC 3'	
pCMB-Tag2B Forward (T3 primer)	5' GCTCGAAATTAACCCTCACTAAAG 3'	1,425 bp
pCMB-Tag2B Reverse (T7 primer)	5' GGTACCTAATACGACTCACTATAGGG 3'	

Table 2: Reagents for PCR insert amplification

Reagent	Amount per reaction
10X Hifi buffer ¹	5 µl
10 mM dNTPs ²	1 µl
50 mM MgSO ₄ ¹	2 µl
Forward Primer (100 mg/ml)	1 µl
Reverse Primer (100 mg/ml)	1 µl
Platinum <i>Taq</i> DNA polymerase High Fidelity (5U/µl) ¹	0.2 µl
Plasmid DNA (1 ng/µl) ³	1 µl
dH ₂ O	38.8 µl
Total volume	50 µl

1. Platinum *Taq* DNA Polymerase High Fidelity Kit (Invitrogen 11304011).
2. 10 mM dNTPs diluted from the 100 mM dNTP set consisting of four deoxynucleotides (dATP, dCTP, dGTP, dTTP) (Invitrogen 10297018).
3. Plasmid DNA: EHMT2/pET28a-LIC plasmid (Addgene 25503) for amplification of G9a-SET cDNA; G9a/pCMV-sport6 plasmid (PlasmID HsCD00327594) for the amplification of the full length G9a, G9a-N-terminal and G9a-ANK cDNAs; CDYL/pCR-BluntII-TOPO plasmid (PlasmID HsCD00347437).

Table 3: PCR setting for PCR cloning of the insert

Temperature	Duration	Cycles
94 °C	2 min	1 cycle
94 °C	30 sec	35 cycles
55 °C	30 sec	
68 °C	42 sec ² , 1:50 min ³ , 3:34 min ⁴	
30 °C	1 sec	1 cycle

1. Eppendorf Mastercycler gradient thermal cycler was used for PCR.
2. PCR setting for amplification of the G9a-SET and G9a-ANK cDNAs.
3. PCR setting for amplification of the G9a-N-terminal cDNAs.
4. PCR setting for amplification of the full length G9a cDNA.

Table 4: Reagents for PCR verification of the insert

Reagent	Amount per reaction
10X PCR buffer ¹	2.5 µl
10 mM dNTPs ²	1 µl
50 mM MgCl ₂ ¹	1.5 µl
Forward Primer (100 mg/ml)	1 µl
Reverse Primer (100 mg/ml)	1 µl
<i>Taq</i> DNA polymerase (5 U/µl) ¹	0.2 µl
Plasmid DNA	1 Colony
dH ₂ O	17.8 µl
Total volume	25 µl

1. *Taq* DNA Polymerase Kit (Invitrogen 18038042).
2. 10 mM dNTPs diluted in dH₂O from the 100 mM dNTP set consisting of four deoxynucleotides (dATP, dCTP, dGTP, dTTP) (Invitrogen 10297018).

Table 5: PCR setting for verification of the insert

Temperature	Time	Cycle
94 °C	5 min	1 cycle
58 °C	1 min	30 cycles
72 °C	1 min	
94 °C	1 min	
55 °C	1 min	1 cycle
72 °C	7 min	
30 °C	1 min	

1. Eppendorf Mastercycler gradient thermal cycler was used for PCR.

Restriction digests of the SET domain and pCS3+MT vector

Restriction digest of the Topo vector containing the SET domain and pCS3+MT empty vector were initially performed with Nco1 restriction enzyme, followed by phenol/chloroform extraction and restriction digest with EcoRI restriction enzyme. Briefly, 2 µg of SET/Topo vector or pCS3+MT, 0.5 µl of the Nco1 (10U/µl, Invitrogen 15421019) restriction enzyme, 5 µl of 10X Buffer K (Invitrogen 15421019), 5 µl of 0.01% BSA (Invitrogen 15421019) were mixed using dH₂O to make a total reaction volume of 50 µl and incubated at 37 °C for 1 hr.

Phenol/chloroform extraction and ethanol precipitation was performed prior to restriction digest with EcoRI restriction enzyme. Briefly, the 50 µl of the SET/Topo vector and pCS3+MT vector from the previous restriction digest with Nco1 was mixed with 1 volume (50 µl) of Tris-saturated phenol and 1 volume (50 µl) of chloroform. The samples were centrifuged at 9,300 x g for 5 min at room temperature. The upper aqueous phase of each sample was transferred to a fresh sterile 1.7 ml tube. An equal volume (50 µl) of chloroform was added to the samples and mixed. The samples were centrifuged at 9,300 x g for 5 min at room temperature and the upper aqueous phase of each sample was transferred to a fresh sterile 1.7 ml tube. The digested SET domain and pCS3+MT plasmids were then precipitated. Briefly, 1/10 the volume (5 µl) of 3 M sodium acetate (pH 5.2) along with twice the volume (10 µl) of 95 % ethanol was added to the samples and incubated at -20 °C for 1 hr. The samples were then centrifuged at 9,300 x g and the supernatant was discarded. The pellet in each sample was washed with 70 % ethanol,

vortexed and centrifuged at 4 °C for 20 min. The ethanol was discarded, and the pellet was resuspended in 3 µl TE buffer.

The samples containing either the SET domain/Topo vector or pCS3+MT vector digested with NcoI restriction enzyme were then used in a second restriction digest using the EcoRI restriction enzyme. SET domain/Topo vector or pCS3+MT [2 µg (3 µl)] was mixed with 0.5 µl of EcoRI restriction enzyme (10U/µl, Invitrogen 15202013), 1 µl of 10X reaction buffer H (Invitrogen 15202013) and dH₂O to make a final reaction volume of 10 µl. They were then incubated at 37 °C for 1 hr.

Gel purification of the restriction digest products and Rapid ligation

After restriction digest with NcoI and EcoRI restriction enzymes, the SET domain insert and the pCS3+MT vector were run on a 1.2 % agarose gel with SYBR Safe DNA Gel Stain (1 in 10,000 dilution) (10,000X, Invitrogen S33102) via gel electrophoresis for 30 min at 100 V. Gel purification of the insert and vector was performed using the Qiagen QIAquick Gel Extraction Kit (28704), following the manufacture's protocol. After gel purification, 1 µl of the insert and the vector were then run on 1 % agarose to verify the gel purified products and estimate their concentration. Rapid ligation of the SET domain insert and pCS3+MT vector was then performed using T4 DNA Ligase Kit (Invitrogen 15224-041), following manufacturer's protocol. A molar ratio of 3:1 insert:vector was used as input. Briefly, the insert and vector were mixed with 1 µl T4 DNA ligase (1U/µl, Invitrogen 15224017), 4 µl of 5X ligase reaction buffer, 1 µl of 10 mM rATP (Promega P1132) with dH₂O making a final reaction volume of 20 µl. The reaction was incubated for 5 min. Subsequently, 2 µl of the ligation reaction was used in transformation of XL1

Blue cells. Briefly, 2 µl of the ligation reaction was gently mixed with 25 µl of XL1 Blue cells and incubated on ice for 20 min. The cells were then heat shocked for 45 sec at 42 °C. The cells were then incubated on ice for 2 min, after which 250 µl of LB was added to each transformation. The cells were then grown for 1 hr on a shaker at 37 °C. The transformed cells (100 µl) were then plated on Amp resistant agar plates and grown overnight at 37 °C. Colonies were selected (18 colonies) and PCR was performed using “G9a-SET Forward” and “G9a-SET Reverse” primers to determine the presence of G9a-SET/pCS3+MT vector using the reagents and settings listed in Tables 4 and 5, respectively. A colony expressing the G9a-SET/pCS3+MT plasmid was inoculated in 200 ml of LB with Amp (50µg/ml) overnight at 37 °C on a shaker. The following day, maxiprep was performed using Qiagen’s Endotoxin-free Maxi Plasmid kit (Qiagen 12362).

E. myc-G9a, myc-G9a-N-terminal, myc-G9a-Ank

Full length G9a, myc-G9a-N-terminal, and myc-G9a-ANK constructs were cloned from G9a/pCMV-sport6 plasmid (PlasmID HsCD00327594, GenBank: BC020970.1), via PCR, into the pCS3+MT vector between EcoRI and Xho1 restriction sites. Initially, the G9a/pCMV-sport6 plasmid was prepared using the Qiagen’s Endotoxin-free Maxi Plasmid kit (Qiagen 12362), following the manufacturer’s protocol. The inserts (full length G9a, G9a-N-terminal and G9a-ANK) were amplified by PRC using Platinum *Taq* DNA polymerase High Fidelity, as previously described, using reagents and PCR setting shown in Tables 2 and 3, respectively. The primers used are listed in Table 1. The PCR products were run on 1.2 % agarose gel SYBR Safe DNA Gel Stain (1 in 10,000 dilution)

(10,000X, Invitrogen S33102) to verify the amplifications. Subsequently, the PCR products were purified using Qiagen QIAquick PCR Purification Kit (Qiagen 28106), following the manufacturer's protocol. The purified PCR products were run on a 1.2 % agarose gel with SYBR Safe DNA Gel Stain (1 in 10,000 dilution) (10,000X, Invitrogen S33102) to estimate the concentration of the PCR products.

Restriction digests and gel purification of the digested products

The purified PCR products (full length G9a, G9a-N-terminal and G9a-ANK cDNAs) were digested with EcoRI and XhoI restriction enzymes. Briefly, for each insert (full length G9a, G9a-N-terminal and G9a-ANK), 2 µg of the PRC product and 2 µg pCS3+MT vector were mixed separately in 1.7 ml sterile tubes with 0.5 µl of EcoRI restriction enzyme (10 U/µl, Invitrogen 15202013), 0.5 µl of XhoI restriction enzyme (10U/µl, Invitrogen 15231012), 2.5 µl of 10X reaction buffer H (Invitrogen 15202013) and dH₂O making a total reaction volume of 25 µl. The samples were incubated at 37 °C for 1 hr. The digested PCR products and pCS3+MT vector were run on a 1.2 % agarose gel with SYBR Safe DNA Gel Stain (1 in 10,000 dilution) (10,000X, Invitrogen S33102) and gel purified using the Qiagen QIAquick Gel Extraction Kit (28704).

Rapid ligation and transformation of XL1 Blue cells

The purified products were run on a 1.2 % agarose gel with SYBR Safe DNA Gel Stain (1 in 10,000 dilution) (10,000X, Invitrogen S33102) to estimate their concentration. Rapid ligation of the inserts (full length G9a, N-terminal G9a, ANK region) and pCS3+MT vector was then performed using T4 DNA ligase kit (Invitrogen 15224-041),

following the manufacturer's protocol. A molar ratio of 3:1 insert:vector was used. Briefly, the insert and vector were mixed with 4 µl of 5X ligase reaction buffer, 1 µl T4 DNA ligase (1U/µl), 1 µl of 10 mM rATP (Promega P1132) with dH₂O making a final reaction volume of 20 µl. The reaction was incubated for 5 min. Subsequently, 2 µl of the ligation reaction was used in transformation of XL1 Blue cells. Briefly, 2 µl of the ligation reaction was gently mixed with 25 µl of XL1 Blue cells and incubated on ice for 20 min. The cells were then heat shocked for 45 sec at 42 °C. The cells were then incubated on ice for 2 min, after which 250 µl of LB was added to each transformation. The cells were then grown for 1 hr on a shaker at 37 °C. The transformed cells (100 µl) were then plated on Amp-resistant agar plates and grown overnight at 37 °C. Colonies were selected (12 colonies for each insert-vector ligation) and PCR was performed using primers listed in Table 1 to determine the presence of G9a/pCS3+MT, G9a-N-terminal/pCS3+MT or G9a-ANK/pCS3+MT plasmid using the reagents and settings listed in Tables 4 and 5, respectively. Colonies expressing the desired plasmid were inoculated in 200 ml of LB with Amp (50µg/ml) overnight at 37 °C on a shaker. The following day, maxiprep was performed using Qiagen's endotoxin-free maxi plasmid kit (Qiagen 12362).

F. pCMV-Tag2B, Flag-CDYL

pCMV-Tag2B is a flag-tagged empty vector. CDYL was cloned from CDYL/pCR-BluntII-TOPO plasmid (PlasmID HsCD00347437, GenBank: BC119682.2) into the pCMV-Tag2B vector at the EcoRI restriction site. Initially, the CDYL/pCR-BluntII-TOPO plasmid was prepared using the Miniprep Kit (Promega A1460), following

the manufacturer's protocol. Nanodrop Spectrometry (Thermo Scientific ND1000) was used, following the manufacturer's protocol, to determine the concentration of the plasmid yielded.

Restriction digests of the CDYL/pCR-BluntII-TOPO plasmid and pCMV-Tag2B empty plasmid and gel purification of digested products

A restriction digest was performed on the CDYL/pCR-BluntII-TOPO and the pCMV-Tag2B plasmids performed using EcoRI restriction enzyme. Briefly, 2 µg of CDYL/pCR-BluntII-TOPO and pCMV-Tag2B plasmids each, in separate tubes, were mixed with 0.5 µl of EcoRI restriction enzyme (10U/µl, Invitrogen 15202013), 5 µl of 10X reaction buffer H (Invitrogen 15202013) and dH₂O, making a total reaction volume of 50 µl, and incubated at 37 °C for 1 hr. After restriction digest, pCMV-Tag2B vector was CIAP treated so the blunt ends do not re-ligate. Briefly, CIAP (Promega M1821) was added directly to digested pCMV-Tag2B plasmid; 5 µl CIAP 10X reaction buffer (Promega M1821), 0.01 U CIAP/pmol of ends and dH₂O to a final volume of 50 µl. It was then incubated at 37 °C for 30 min. A second aliquot of diluted CIAP (equivalent to the amount used initially) was added and incubated at 37 °C for an additional 30 min. The digested CDYL/pCR-BluntII-TOPO and pCMV-Tag2B plasmids were run on a 1.2 % agarose gel with SYBR Safe DNA Gel Stain (1 in 10,000 dilution) (10,000X, Invitrogen S33102) for 30 min at 100 V and gel purified using the Qiagen QIAquick Gel Extraction Kit (Qiagen 28704), following the manufacturer's protocol.

Rapid ligation and transformation of XL1 Blue cells

The purified products were run on a 1.2 % agarose gel with SYBR Safe DNA Gel Stain (1 in 10,000 dilution) (10,000X, Invitrogen S33102) to estimate their concentrations. Rapid ligation of the CDYL insert and pCMV-Tag2B vector was performed using T4 DNA Ligase Kit (Invitrogen 15224-041), following manufacturer's protocol. A molar ratio of 3:1 insert:vector was used. Briefly, the insert and vector were mixed with 5 µl of 5X ligase reaction buffer, 1 µl T4 DNA ligase (1U/µl), 1 µl of 10 mM rATP (Promega P1132), with dH₂O to make a final reaction volume of 25 µl. The reaction was incubated for 5 min. Subsequently, 2 µl of the ligation was used in transformation of XL1 Blue cells. Briefly, 2 µl of the ligation was gently mixed with 25 µl of XL1 Blue cells and incubated on ice for 20 min. The cells were then heat shocked for 45 sec at 42 °C. The cells were then incubated on ice for 2 min, after which 250 µl of LB was added to each transformation. The cells were then grown for 1 hr on a shaker at 37 °C. The transformed cells (100 µl) were then plated on Amp-resistant agar plates and grown overnight at 37 °C. Colonies were selected (18 colonies) and PCR was performed using primers listed in Table 1 (pCVM-Tag2B vector primers) to determine the presence of CDYL/pCVM-Tag2B plasmid using the reagents and settings listed in Tables 4 and 5, respectively. Colonies expressing the CDYL/pCVM-Tag2B plasmid was inoculated in 200 ml of LB with Amp (50µg/ml) overnight at 37 °C on a shaker. The following day, the CDYL/pCVM-Tag2B plasmid was prepared using using Qiagen's endotoxin-free maxi plasmid kit (Qiagen 12362).

2.11 Plasmid preparation

Qiagen's Endotoxin-free Maxi Plasmid kit (Qiagen 12362) was used to prepare all plasmids. The sequences/mutations were confirmed by automated dideoxynucleotide sequencing of both forward and reverse nucleotide strands, by the Centre for Applied Genomics, the Hospital for Sick Children, Toronto, Canada.

Table 6: Plasmids

Plasmid Name	Description
GST-MIER1 β	GST-tagged full length MIER1 β . MIER1 β cloned into pGEX-4T-1 vector.
GST-aa1-283	GST-tagged MIER1 β construct protein containing amino acids 1-283 (cloned into pGEX-4T-1 vector). It contains the ELM2 domain, but lacks the SANT domain.
GST-aa287-512	GST-tagged MIER1 β construct containing amino acids 287-512 (cloned into pGEX-4T-1 vector). It contains only the SANT domain.
GST-aa287-357	GST-tagged MIER1 β construct containing amino acids 281-357 (cloned into pGEX-4T-1 vector). It contains the SANT domain and the β specific C terminus.
GST-SANTmutant	GST-tagged full length MIER1 β with three point mutations at ³²⁶ FYY \rightarrow AAA
pCS3+MT	Myc-tagged empty vector.
Myc-MIER1 α	Myc-tagged full length MIER1 α . MIER1 α cloned into pCS3+MT vector.
myc-G9a	Myc-tagged full length G9a protein (cloned into pCS3+MT vector).
myc-G9a-N-terminal	Myc-tagged G9a construct containing amino acids 21-628 (cloned into pCS3+MT vector).
myc-G9a-Ank	Myc-tagged G9a construct containing amino acids 649-879 (cloned into pCS3+MT vector). It contains only the ankyrin repeats of G9a protein.
myc-G9a-SET	Myc-tagged G9a construct containing amino acids 913-1142 (cloned into pCS3+MT vector). It contains only the SET domain of G9a protein.
pCMV-Tag2B	Flag-tagged empty vector.
Flag-CDYL	Flag-tagged full length CDYL protein. It is cloned into flag-tagged pCMV-Tag2B empty vector.

2.2 Cell culture

HEK-293 cell line was purchased from the American Tissue Culture Collection and cultured in DMEM (Gibco 11965-092) supplemented with 10 % serum (7.5 % newborn calf serum (Gibco 16010-142), 2.5 % fetal calf serum (Gibco 16000-044), 1 mM sodium pyruvate (Gibco 11360-070) and 0.5 % penicillin/streptomycin in a humidified 37 °C incubator with 5 % CO₂.

2.3 Transfection of HEK-293 cells

Cells were seeded at a density of 5×10^5 cells/ well in 6 well plates and grown for 24 hr in DMEM without antibiotics. Following the 24 hr period, cells were transfected using the Mirus TransIt-LT1 (Fisher Scientific, MIR 2300) reagent according to manufacturer's protocol. Briefly, 250 µl of serum-free, antibiotic-free Opti-MEM (Gibco 51985-034) was placed in a sterile tube per transfection. Mirus TransIT-LT1 reagent was warmed to room temperature and then added to the 250 µl media, using a 3:1 µl Mirus TransIT-LT1 reagent to µg DNA ratio. The solution was gently mixed and plasmid DNA was added to the diluted Mirus TransIT-LT1 reagent. The solution was gently mixed again and incubated at room temperature for 15 min, following which the Mirus TransIT-LT1 Reagent – DNA complex was added drop-wise to the cells. The dish was gently rocked back and forth and from side to side for even distribution of the Mirus TransIT-LT1 Reagent – DNA complexes. The cells were then incubated at 37 °C and 5% CO₂ for 48 hr, upon which they were subject to protein extraction and used for immunoprecipitation.

2.4 GST pull-down assay

2.4.1 GST-fusion protein production

Purified GST protein was kindly provided by Jarratt Rose. GST-tagged MIER1 constructs either lacking or containing the SANT domain (aa1-283, aa287-512, aa287-357) and full length MIER1 β were prepared by Corinne Mercer, and GST- SANT mutant (³²⁶FYY \rightarrow AAA) was made by Julia Ferguson. The GST-SANT mutant is full length GST-MIER1 β protein with 3 point mutations at amino acids 326-328, where the phenylalanine and the following two tyrosine residues are mutated to alanine (³²⁶FYY \rightarrow AAA).

Briefly, GST and GST-tagged MIER1 constructs were prepared as follows. The pGEX vectors expressing these proteins were transformed into Rosetta™ 2(DE3)pLysS Competent Cells (EMD Millipore 71403-3). The plasmid DNA (10 ng) was gently mixed with 25 μ l of Rosetta cells and incubated on ice for 20 min, following which the cells were heat shocked for 45 sec at 42 °C. The cells were then incubated on ice for 2 min, after which 250 μ l of LB was added to each transformation. The Rosetta cells were then grown for 1 hr on a shaker at 37 °C. The transformed cells (100 μ l) were then plated on Amp-resistant agar plates and grown overnight at 37 °C. From the transformed cells, a 5 ml culture of LB with Amp (50 μ g/ml) was inoculated and grown overnight on a shaker at 37 °C. The following day, 250 ml of LB with Amp (50 μ g/ml) was inoculated with 1 ml of the overnight culture. It was grown at 37 °C on a shaker to an optical density (OD₆₀₀) of 0.8-1.0, at which the culture was induced to make the GST or GST-fused

protein by adding 25 μ l of 1M Isopropyl β -D-1-thiogalactopyranoside (IPTG). The culture was grown on a shaker overnight at room temperature. The culture was then transferred aseptically to a 250 ml Nalgene bottle and centrifuged at 1500 x g at 4 °C for 10 min. The supernatant was drained and the pellet was frozen and thawed to help burst the cells open during sonication. The pellet was resuspended in ice cold 5 ml of 1xPBS and transferred to a 50 ml Falcon tube. 50 μ l of 0.2 M phenylmethylsulfonyl fluoride (PMSF) (protease inhibitor) was added and the sample was sonicated on ice; 2 pulses, 15 sec each with 30 sec pause between pulses, 70 % amplitude (Fisher Scientific Sonic Dismembrator Model 500). The sample was then transferred to a 30 ml Corex tube and 500 μ l of 10 % TritonX-100/1xPBS was added. The sample was then centrifuged at 9,300 x g at 4 °C for 15 min and the aqueous layer was transferred to clean 1.7 ml tubes in 1 ml aliquots and stored at -70 °C for future applications.

Prior to use, the expression of GST fusion proteins was verified. For each protein, 50 μ l of a 50 % slurry of Glutathione Sepharose 4B beads was washed twice with GST pull-down buffer [20 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 10 % glycerol, 0.2 % NP-40 (in PBS)] in a 1.7 ml tube; 1 ml of GST pull-down buffer was added to the beads, inverted to mix and centrifuged for 30 sec at 500 x g. The supernatant was aspirated off. The beads were made into a 50 % slurry using GST pull-down buffer. Subsequently, 50 μ l of the beads were then added to GST pull-down buffer containing the soluble protein being tested, along with 5 μ l of 0.2 M PMSF in a final 1 ml volume. Three different volumes of the GST-fused proteins were tested; 50 μ l, 100 μ l and 200 μ l. GST pull-down buffer was added to make a total volume of 1 ml for each sample. The samples

containing the beads were then rotated at 4 °C for 1 hr, after which they were washed twice with 1 ml of GST pull-down buffer, and once with 1ml of 150 mM NaCl. The supernatant was aspirated and 30 µl of 1.5 x SSB [SDS sample buffer (0.125 M Tris-HCl pH 6.8, 2 % SDS, 5 % β-mercaptoethanol, 20 % glycerol)] + bromophenol blue (dye) was added to the beads. The samples were then boiled for 4 min and placed on ice for 2 min. The samples were centrifuged at 500 x g for 30 sec and the aqueous layer was loaded on a 7.5 % SDS-polyacrylamide gel (BioRad 456-1024). Bovine serum albumin (BSA RIA grade 1) (Sigma A7888) was also loaded on the gel for determining the approximate amount of GST proteins via comparison to known BSA levels. After the gel was run, it was removed from the glass plate and placed in Coomassie Blue Stain (0.025 % Coomassie brilliant blue (BioRad 161-0400), 45 % straight methanol, 10 % glacial acetic acid) for 60 min. It was then de-stained overnight with de-stain buffer (30 % straight methanol, 10% glacial acetic acid) to remove the excess dye. The gel was then dried on Whatman paper for approximately 1 hr at 80 °C under vacuum.

2.4.2 GST pull-down assay

GST pull-down assay was performed using Glutathione Sepharose 4B beads (Qiagen) and GST-tagged MIER1β constructs. For each condition, 50 µl of the 50 % slurry of Glutathione sepharose beads were washed with 1 ml of GST pull-down assay buffer, followed by a 1 min rotation at room temperature, and centrifugation at 500 x g for 30 sec. The supernatant was then aspirated off and the procedure repeated twice. Subsequently, beads were resuspended in 1 ml of GST pull-down buffer and were incubated with the GST-fused MIER1 proteins (154 nM GST, 23 nM aa1-283, 10 nM

aa287-512, 59 nM aa287-357, 10 nM SANTmutant and 12 nM full length MIER1 β ; diluted in GST pull down buffer) for one hr at 4 °C on a rotator. Following the 1 hr incubation period, the beads were washed 3 times as previously described. Beads were then resuspended in 1 ml interaction buffer (GST pull-down buffer with 1mg/ml BSA and 1mM PMSF) along with 10 nM recombinant G9a protein (diluted in GST pull-down buffer) (Origene, TP319200), and rotated for another hr at 4 °C. Beads were then washed 3 times with the GST pull-down buffer and twice with 1 ml of 150 mM NaCl. Subsequently, the beads were resuspended in 30 μ l of 1.5 x SSB + bromophenol blue (dye). Following which, the samples were boiled for 3 min, the supernatant was collected and loaded on a 7.5 % SDS-polyacrylamide gel (BioRad 456-1024) and western blot analysis was performed. Trans-blot Turbo System was used to transfer the gels onto PVDF membranes (BioRad-Trans-blot Turbo Transfer pack 0170-4156) at voltage limit of 25 V and constant current of 2.5 A for 10 min. G9a antibody (1 in 1000 dilution of 1 μ g/ μ l stock, Millipore 07-441) was used to probe for G9a and GST antibody (1 in 5000 dilution of 0.5 μ g/ μ l stock, Genscript A00865) was used to detect the level of GST-fused proteins used. Table 7 lists all the primary antibodies used. Densitometric analysis was carried out using ImageJ software. Densitometric measurements of the amount of G9a bound were divided by the densitometric measurements of the amount of GST/GST-MIER1 β constructs to obtain relative amount of G9a interaction with each input.

2.5 Immunoprecipitation

2.5.1 Cell lysate preparation

For *in vivo* immunoprecipitation (IP) assay between myc-G9a constructs and MIER1 α , HEK-293 cells were seeded at a density of 5×10^5 cells/well in a 6-well plate and grown in DMEM for approximately 24 hr. Subsequently, cells were transfected using Mirus TransIT-LT1 Reagent as described on page 54, with 0.5 μ g of MIER1 α and myc-FL-G9a or myc-G9a-N-terminal or myc-G9a-ANK or myc-G9a-SET-expressing plasmids; 0.5 μ g MIER1 α -expressing plasmid and pCS3+MT empty vector; 0.5 μ g myc-FL-G9a-expressing plasmid and pcDNA3.1 empty vector. For *in vivo* immunoprecipitation (IP) assay between MIER1 α , flag-tagged CDYL, and myc-tagged G9a, cells were transfected with 0.5 μ g of MIER1 α , myc-G9a and flag-CDYL plasmids; 0.5 μ g of MIER1 α , myc-G9a plasmids and pCMV-tag2b vector; 0.5 μ g of MIER1 α plasmid, pCS3+MT vector and flag-CDYL plasmid; 0.5 μ g pcDNA3.1 vector, myc-G9A and flag-CDYL plasmids; MIER1 α plasmid, pCS3+MT empty vector and pCMV+tag2b empty vector. After 48 hr, cells were lysed as follows. Each well was washed with 2ml of 1x PBS at room temperature, and 500 μ l of 1x Triton lysis buffer diluted from a 10x stock (0.1 M Tris pH 7.5, 0.1 M EDTA, 0.2 % sodium-azide, 10 % Triton X-100 in dH₂O) containing 100x PI and 1 mM PMSF; two wells were then combined for each condition. A third well was extracted with 500 μ l of 1.5 x SSB. After 1xTriton or SSB were added, the cells were scraped with cell lifters (Fisher Scientific 08100240) and incubated for 20 min on ice. Cell lysates were then pulled through a 1 ml syringe 30 times

to shear the DNA, following which they were transferred to labeled, chilled 1.7 ml tubes. The cell lysates were then centrifuged at 12,000 x g for 10 min at 4°C and the supernatant was transferred to another 1.7 ml tube containing protein A agarose beads (Pierce 20333), prepared as described below.

2.5.2 Co-immunoprecipitation

Protein A agarose beads (Pierce 20333) were washed with 1x Triton medium, following a 1 min rotation at room temperature, and centrifugation at 500 x g for 30 seconds. This was repeated twice, and the final pellet was made into a 50 % slurry of the final pellet with 1x Triton buffer. For each IP condition, 50 µl of the 50 % slurry was used. For IP between the G9a constructs and MIER1α and for IP between MIER1α, flag-tagged CDYL and myc-tagged G9a, MIER1α antibody was used to pre-bind to the beads (1 in 100 dilution). MIER1α antibody was added to the beads along with 990 µl 1 x Triton medium and 10 µl PI, and incubated on a rotator for 1 hr at 4 °C; preimmune serum was used as a negative control. For the reverse IP between myc-G9a constructs and MIER1α, 9E10 antibody was pre-bound to the beads (1 in 100 dilution). The beads were then centrifuged for 30 sec at 0.5 x g and washed twice as described above. After the cell lysates were prepared as previously described, they were loaded onto the beads and incubated overnight at 4 °C. The samples were washed twice with 1 x Triton medium, and once with 150 mM NaCl. The beads were centrifuged for 1 min at 500 x g and resuspended in 30 µl 1.5 x SSB + bromophenol blue dye and boiled for 4 min. The beads

were then centrifuged again (as described above) and the supernatant was loaded on a 10 % SDS-PAGE gel (BioRad 456-1034).

2.5.3 Western blot analysis

SDS-PAGE gels (10%) (BioRad 456-1034) were run at 15 mA for approximately 1.5 hr and were then transferred onto PVDF membranes (BioRad-Trans-blot Turbo Transfer pack 171-4146) via Trans-blot Turbo system (BioRad) at voltage limit of 25 V and constant current of 2.5 A for 10 min. The membrane was then blocked in 5 % skim milk dissolved in 1x Tris-Buffered Saline-Tween 20 (TBST) (20 mM Tris pH 7.6, 150 mM NaCl, 0.1 % Tween-20, and dH₂O) at room temperature for 1 hr. For myc-G9a constructs and MIER1 α IPs, 9E10 antibody (1 in 1000 dilution in 5 % skim milk/1xTBST) was used to probe for myc-tagged G9a and the various myc-tagged G9a constructs, and pan-MIER1 antibody (1 in 1000 dilution) was used to probe for MIER1 α . For IP between MIER1 α , myc-tagged G9a, and flag-tagged CDYL, 9E10 antibody was used to detect myc-tagged G9a and flag antibody (1 in 2000 dilution, Sigma F3165) was used to probe for flag-tagged CDYL. For western blot detection of EZH2, EZH2 antibody (1 in 1000, Cell Signaling 5246S) was used. β -actin antibody (1 in 5000 dilution, Sigma A5441) was used to probe for β -actin. Table 7 on page 70 lists the antibodies used for western blot analysis (the specificities of the antibodies were confirmed prior to use). All antibodies were in 5 % skim milk/1x TBST. The western blots were then incubated on a shaker for 3 hr at room temperature, with the exception of β -actin antibody which was incubated for 1 hr at room temperature. Subsequently, the membranes were washed

several times for 1 hr with a total of 1 liter 1 x TBST. The western blots were then probed with 1 in 5000 dilution of goat anti-mouse conjugated to horseradish peroxidase secondary antibody (0.8 mg/ml stock, Jackson ImmunoResearch 115035174) to detect anti-mouse antibodies, or 1 in 2500 dilution of donkey anti-rabbit conjugated to horseradish peroxidase secondary antibody (GE Healthcare Life Sciences NA934V) to detect EZH2. Table 8 lists the secondary antibodies used. Both secondary antibodies were diluted in 5 % skim milk /1x TBS-T, in which the membranes were incubated for 1 hr. Subsequently, the membranes were washed several times with 1xTBST for 1 hr at room temperature with a total volume of 1 L. The blots were then analyzed using the Clarity Blotting system (BioRad 170-5061)/ ECL prime western blotting system (GE Healthcare RPN2232) and visualized on ECL hyperfilm (Amersham 45001508). Densitometric analysis on CDYL expression was carried out using ImageJ software.

2.6 Methylation assay using recombinant G9a protein

2.6.1 Methylation assay using H3(21-44) peptide as the substrate

Recombinant purified full length G9a (Origene TP319200), and SET domain (Active Motif 31425) were used in this assay. Recombinant SET domain protein contains only the SET domain of the G9a protein with its adjacent cysteine rich regions. It includes amino acids 913-1193 of the G9a protein, which is responsible for the methylation activity. Methylation assay buffer made of 50 mM Tris-HCl pH 9, 1 mM DTT, 0.01 % BSA and 1 mM PMSF was used to dilute the proteins and inhibitors used in the assay. Recombinant full length G9a, or G9a-SET domain were added to microfuge tubes at 2.5

nM and 25 nM, and 2.5 nM and 12.5 nM concentrations, respectively, with and without the inhibitors. The samples were incubated without inhibitors or with 200 μ M S-adenosyl homocysteine (SAH) (Sigma A9384) or with 17 μ M BIX-01294 (Sigma B9311) for 10 min at 37 °C. SAH is a common methyltransferase inhibitor (acts via product inhibition), and BIX-01294 is a G9a-specific inhibitor. After the 10 min incubation period, the samples were incubated with 20 μ M S-adenosyl-methionine (SAM) (Sigma A4377) for another 10 min. Subsequently, 17 μ M H3(21-44) biotinylated peptide (Anaspec 64440) diluted in the methylation assay buffer was added to each sample and incubated for 40 min at 37 °C. Each sample had a total reaction volume of 100 μ l. SAM was added to the reaction prior to the peptide substrate since a study done by Patnaik *et al.* (2004) showed that this causes G9a to have higher methylation activity [39].

Streptavidin coated, high-capacity plates (Pierce Thermo Scientific 15501) were used in this assay to capture the biotinylated peptide as avidin and biotin have high binding affinity towards each other. After the 1 hr total reaction time, streptavidin 8-well strips were washed twice with 1x TBST (20 mM Tris pH 7.6, 150 mM NaCl, 0.1 % Tween-20, and dH₂O). The samples were then loaded onto the wells, and incubated for 1 hr at 4 °C, allowing the biotinylated H3(21-44) peptide to bind to the avidin immobilized on the wells. Following the incubation, the wells were washed three times with 1xTBST and blocked with 150 μ l of superbloc T20 (TBS) blocking buffer (Thermo Scientific 37536) for 30 min. After the blocking period, polyclonal histone H3K27me1 (1 μ g/ μ l stock, Active Motif 39889) and H3K27me2 (1 μ g/ μ l stock, Active Motif 39919) antibodies in 1xPBS (1 in 200 dilution) were added to the wells. The primary antibody

was incubated for 1 hr at room temperature, following which the wells were washed three times with 1xTBST. Anti-rabbit conjugated to horseradish peroxidase secondary antibody in 1x PBS (1 in 2000) was then added to the wells with a 1 hr incubation period. Subsequently, the wells were washed once again 3 times with 1x TBST, and 150 μ l of 1-Step ABTS Solution (Thermo Scientific 37615) was added to each well, and incubated for 30 min. The reaction was quenched with 100 μ l of stop solution (1 % SDS with 0.02 % sodium azide in dH₂O). The 8-well strips were then read on a plate reader (BioRad Microplate Reader 3550) at λ = 415 nm. Each condition in each experiment was performed in duplicate. The data were normalized and visualised as histogram.

2.6.2 Methylation assay using recombinant histone 3 as the substrate

To further elucidate the role of G9a in H3K27 methylation, the activity of recombinant SET domain (Active Motif 31425) was tested on whole, recombinant histone 3.1 (New England Biolabs M2503S), and methylation was detected via western blot analysis. This assay was performed with 44 ng SET domain, 0.75 μ g recombinant histone 3.1, 5 μ M SAM, and 150 μ M SAH, in methylation assay buffer (50 mM Tris-HCl pH 9, 1mM DTT, 0.01 % BSA, 1 mM PMSF) with a total reaction volume of 30 μ l. Initially, the SET domain was diluted in the assay buffer and incubated with 150 μ M SAH or equal amount of buffer for 10 min at 37 °C. Subsequently, 5 μ M SAM was added to each sample, and incubated for 10 min at 37 °C. The substrate, 0.75 μ g recombinant histone H3.1, was then added to each sample and the reaction was incubated for additional 1 hr 40 min at 37 °C. After the incubation period, 10 μ l of 4x SSB (SDS sample buffer [200 mM Tris HCl, 8 % SDS, 20 % β -mercaptoethanol, 40 % glycerol]) + bromophenol blue

(dye) was added to each sample. The samples were then boiled and loaded on a 4-20 % gradient SDS-polyacrylamide gel (BioRad 456-1094). The gels were then transferred onto nitrocellulose membrane (BioRad 162-0212) via Trans-blot Turbo System (BioRad) at voltage limit of 25 V and constant current of 1.3 A for 5 min. Western blot analysis was performed as described earlier. H3K27me2 antibody (1 in 1000 dilution of 1 µg/µl stock, Active Motif 39919) was used to detect H3K27 methylation, and H3K9me2 antibody (1 in 1000 dilution of 1 µg/µl stock, Millipore 07-441) was used to probe for H3K9 methylation. H3K9 methylation was tested as a positive control.

The efficacy of G9a specific inhibitor (BIX-01294, Sigma B9311), EZH2 specific inhibitor (GSK-343, Sigma SML0766) and SAH (Sigma A9384) on recombinant SET protein was tested in subsequent experiment as well. This was performed to determine if BIX and GSK-343 inhibit G9a at the selected concentrations, which may be used in future experiments. Only H3K9 methylation was tested as previous experiment showed no detectable H3K27 methylation by G9a. The assay was performed as previously described, using methylation assay buffer (50 mM Tris-HCl pH 9, 1 mM DTT, 1 mM PMSF), 0.44 ng G9a-SET protein, 0.75 µg recombinant histone 3.1, 5 µM SAM, 150 µM SAH, 170 µM BIX-01294, 40 nM GSK-343 and 4 µM GSK-343. All the reagents used were initially diluted in the methylation assay buffer. Following the methylation assay, western blot analysis was performed as described above. H3K9me2 antibody (1 in 1000 dilution of 1 µg/µl stock, Millipore, 07-441) was used to detect methylation.

2.7 Methylation assay using GST-MIER1 β protein and recombinant H3(21-44) peptide as the substrate

To determine the role of MIER1 on H3K27 methylation, GST-MIER1 β was used, along with GST protein as a negative control. GST protein was kindly provided by Jarratt Rose, and full length GST-MIER1 β was prepared by Corinne Mercer as previously described (pages 55-57). For this assay, a GST pull-down was initially performed as previously described (pages 57-58), with minor alterations. Briefly, for each condition 154 nM GST or 12 nM GST-MIER1 protein was added to 1 ml of GST pull-down buffer containing 50 μ l of 50 % slurry of glutathione sepharose beads, and incubated for 1 hr at 4 °C on a rotator. Following the 1 hr incubation period, the beads were washed 3 times as previously described (page 57). The GST and GST-MIER1 protein bound to the beads were subsequently used in the methylation assay previously described (pages 62-64). The beads bound with GST or GST-MIER1 β were incubated with 200 μ M SAH, 17 μ M BIX-01294, or with equal volume methylation assay buffer for 10 min at 37 °C on a rotator. Subsequently, the samples were incubated with 20 μ M SAM for another 10 min. Inhibitors, SAH and BIX-01294, and the methyl donor, SAM, were diluted in the methylation assay buffer (50 mM Tris-HCl pH 9, 1 mM DTT, 1 mM PMSF). Following the 10 min incubation period 17 μ M H3(21-44) peptide, also diluted in the methylation assay buffer, was added to each sample and incubate for additional 40 min at 37 °C. Each sample had a total reaction volume of 100 μ l and a total reaction time of 1 hr. After the 1 hr period, the beads were centrifuged at 500 x g for 30 seconds. The supernatant containing the reaction mixture was collected and added to the streptavidin 8-well strips.

The detection of the H3K27 methylation followed the same procedure as with methylation assay using recombinant G9a previously described (pages 62-64). The wells were read by the plate reader and the results were presented as histograms.

2.8 Methylation assay using MIER1 α complex from HEK293 cells

HEK-293 cells used to determine if MIER1 α immunoprecipitated from cell extracts is able to methylate H3K27. As with the immunoprecipitation assay previously described, HEK-293 cells were seeded at a density of 5×10^5 cells/well in a 6-well plate and grown in DMEM for approximately 24 hr. Cells were then transfected using Mirus TransIT-LT1 Reagent with 0.5 μ g of pCS3+MT (empty vector), or myc-MIER1 α expressing plasmids. After 48 hr, cells were lysed as follows. Each well was washed with 2 ml of 1x PBS at room temperature, and 500 μ l of 1 x Triton lysis buffer diluted from a 10x stock (0.1 M Tris pH 7.5, 0.1 M EDTA, 0.2 % sodium-azide, 10 % Triton X-100 in dH₂O) containing 100x PI and 1 mM PMSF; two wells were then combined for each condition. After 1x Triton was added, the cells were scraped with cell lifters (Fisher Scientific) and incubated for 20 min on ice. Cell lysates were then pulled through a 1 ml syringe 30 times to shear the DNA, following which they were transferred to labeled, chilled 1.7 ml tubes. The cell lysates were then centrifuged at 12,000 x g for 10 min at 4 °C and the supernatant was transferred to another 1.7 ml tube containing protein A agarose beads pre-bound with 9E10 antibody (1 in 100 dilution).

Protein A agarose beads were prepared as follows. Protein A beads (Pierce) were washed with 1x Triton medium, following a 1 min rotation at room temperature, and centrifugation at 0.5 x g for 30 sec. This was repeated twice, and the final pellet was made

into a 50 % slurry of the final pellet with 1x Triton buffer. For each IP condition, 50 μ l of the 50% slurry was used. MIER1 α antibody was used to pre-bind to the beads (1 in 100 dilution). MIER1 α antibody was added to the beads along with 990 μ l 1 x Triton medium and 10 μ l 100x PI, and incubated on rotator for 1 hr at 4 °C; preimmune serum was used as a negative control. The beads were then centrifuged for 30 sec at 0.5 x g and washed twice as described above. After the cell lysates were prepared as described above, they were loaded onto the beads and incubated overnight at 4 °C. The following day, the samples were washed twice with 1 x Triton medium, and once with 150 mM NaCl. The beads were then resuspended in the methylation assay buffer and used in the methylation assay as follows.

The immunoprecipitated cell lysates were incubated with 200 μ M SAH, 17 μ M BIX-01294, or with equal volume methylation assay buffer for 10 min at 37 °C on a rotator. Subsequently, the samples were incubated with 20 μ M SAM for another 10 min. Inhibitors, SAH and BIX-01294, and the methyl donor, SAM, were diluted in the methylation assay buffer (50 mM Tris-HCl pH 9, 1 mM DTT, 1 mM PMSF). Following the 10 min incubation period 17 μ M H3(21-44) peptide, also diluted in the methylation assay buffer, was added to each sample and incubate for additional 40 min at 37 °C. Each sample had a total reaction volume of 100 μ l and a total reaction time of 1 hr. After the 1 hr period, the beads were centrifugation at 0.5 x g for 30 sec. The supernatant containing the reaction mixture was collected and added to the streptavidin 8-well strips. The detection of H3K27 methylation followed the same procedure as with methylation assay

using recombinant G9a previously described (pages 62-64). The wells were read by the plate reader and the results were presented as histograms.

Table 7: Primary antibodies

Name	Description
9E10	Mouse monoclonal anti-myc antibody
β -actin	Mouse monoclonal anti- β -actin antibody purchased from Sigma (A5441).
EZH2	Rabbit monoclonal anti-EZH2 antibody purchased from Cell Signaling (5246S).
GST	Mouse monoclonal anti-GST antibody purchased from Genscript (A00865).
G9a	Rabbit polyclonal anti-G9a antibody purchased from Millipore (07-551).
H3K9me2 antibody	Rabbit polyclonal anti-dimethyl-histone H3 (Lys9) antibody purchased from Millipore (07-441).
H3K27me1 antibody	Rabbit polyclonal anti-monomethyl-histone H3 (Lys27) antibody purchased from Active Motif (39889).
H3K21me2 antibody	Rabbit polyclonal anti-monomethyl-histone H3 (Lys27) antibody purchased from Active Motif (39919).
Histone H3.1 antibody	Rabbit monoclonal anti-histone H3.1 antibody purchased from Cell Signalling (4499S).
Flag	Mouse monoclonal anti-Flag M2 antibody purchased from Sigma (F3165). Immunogen sequence: DYKDDDK.
pan-MIER1	Mouse monoclonal MIER1 antibody. It recognizes both α and β MIER1 isoforms. Prepared by immunizing rabbits with $^{16}\text{CSDDHEFGPSTDMLVHD}^{32}$ synthetic peptide of the predicted MIER1 protein sequence [101]
MIER1 α	In house rabbit polyclonal MIER1 α specific antibody. Prepared by immunizing rabbits with $^{413}\text{CQMLLPVHFSAISSR}^{426}$ synthetic peptide of the predicted MIER1 protein sequence [101].

Table 8: Secondary antibodies

Name	Description
Donkey anti-rabbit	Donkey anti-rabbit conjugated to horseradish peroxidase secondary antibody purchased from GE Healthcare Life Sciences (NA934V).
Goat anti-mouse	Light chain specific goat anti-mouse conjugated to horseradish peroxidase secondary antibody purchased from Jackson ImmunoResearch (115035174).

Chapter 3: Results

3.1 *Interaction between MIER1 and G9a*

3.1.1 *In vitro* interaction between MIER1 β and G9a

To determine whether MIER1 β and G9a interact *in vitro*, GST pull-down assays were performed. Moreover, GST-tagged deletion constructs of MIER1 β (Fig. 6A) were used in the GST pull-down to further determine, if the two proteins interact, which region of MIER1 is involved in this interaction. GST-fusion constructs either lacking or containing the SANT domain [GST, aa1-283, aa287-512, aa287-357, SANTmutant(³²⁶FYY-AAA) and full length MIER1 β] were used to determine which region in MIER1 plays an important role in the recruitment of G9a. Recombinant full length G9a (Origene) was used for this assay. The relative amount of GST-fused proteins used in the assay is shown by Western blot analysis (Fig. 6D).

The GST pull-down assay showed that MIER1 β interacts with G9a *in vitro* as evident in Fig. 6B-C. Moreover, the region of MIER1 that is responsible for this interaction contains the SANT domain. The construct that lacks the SANT domain, aa1-283, does not interact with G9a (Fig. 6B), further confirming the role of SANT domain in MIER1-G9a interaction. Nonetheless, the SANT mutant (which contains the FYY to AAA amino acid change at aa326-328) also interacts with G9a (Fig. 6B), which suggests that these conserved amino acids are not responsible for MIER1-G9a interaction. Furthermore, there is some interaction seen with GST alone; however, this could be

because of the relatively higher amount of GST protein used (Fig. 6D), causing nonspecific binding. When the amount of G9a bound is normalised to the amount of GST/GST-MIER1 β constructs, there is greater amount of G9a binding to GST-MIER1 β than to GST protein (Fig. 6C).

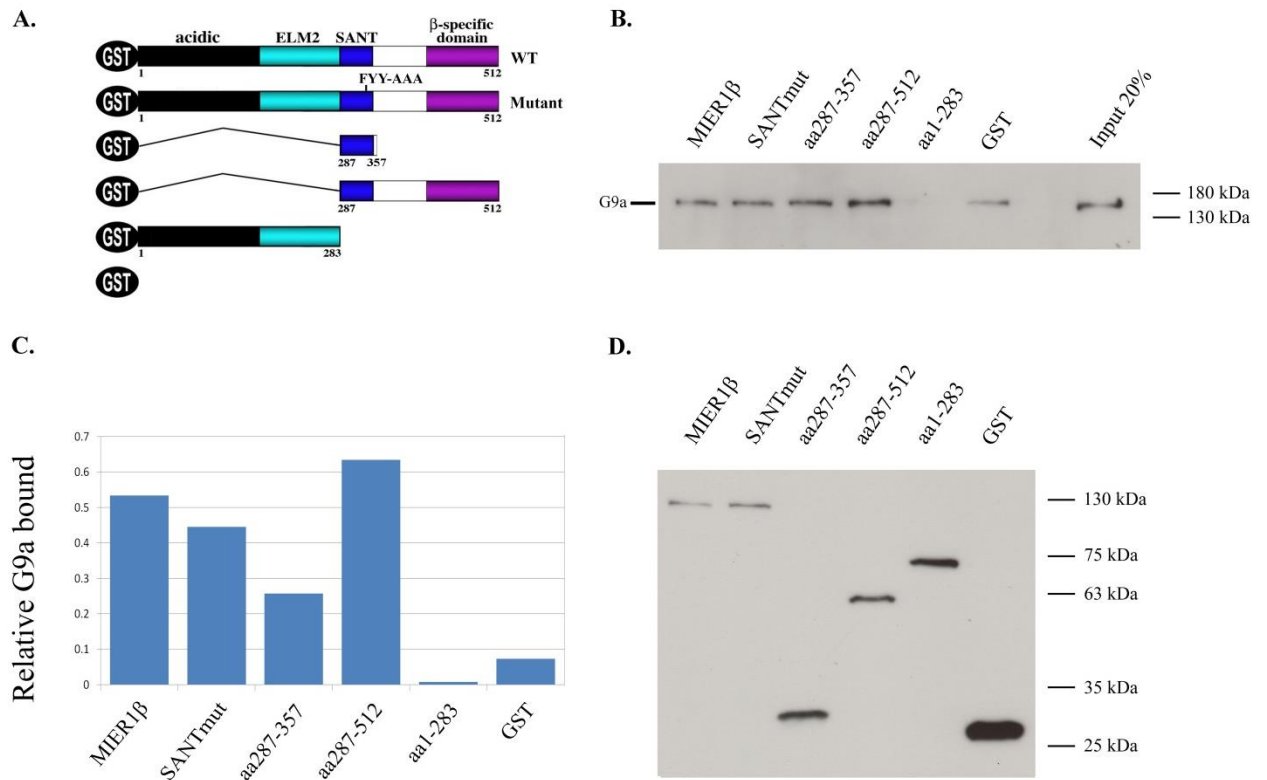


Figure 6: GST pull-down showing the interaction between G9a and GST-MIER1 β fusion constructs *in vitro*

Interaction between GST-fusion proteins [GST, aa1-283, aa287-512, aa287-357, SANTmutant(³²⁶FYY-AAA) and MIER1 β] and purified recombinant G9a (Origene). GST-fusion proteins were prepared and GST pull-down assay was performed three times as described in Materials and Methods section. The GST pull-down assay was repeated three times. **A)** Schematic representation of the GST-MIER1 β constructs used in the assay. **B)** The amount of G9a binding to GST and GST-MIER1 constructs. G9a antibody was used to probe for the recombinant G9a protein. **C)** The relative amount of G9a bound as calculated using ImageJ densitometry analysis of the blots (level of G9a bound/amount of GST-MIER1 β construct). **D)** The levels of GST-MIER1 β constructs used in the assay, detected using GST antibody.

3.1.2 *In vivo* interaction between MIER1 α and G9a

To determine whether MIER1 α and G9a interact *in vivo*, HEK-293 cells were transiently transfected with 0.5 μ g of MIER1 α and full length G9a (myc-FL-G9a) or the N terminal (myc-G9a-N-terminal; aa 21-628) or the ankyrin repeats (myc-G9a-ANK; aa 649 – 879) or the SET domain of G9a (myc-G9a-SET; aa 913-1142)-expressing plasmids; 0.5 μ g MIER1 α -expressing plasmid and pCS3+MT empty vector; 0.5 μ g myc-FL-G9a-expressing plasmid and pcDNA3.1 empty vector. Fig. 7A shows the myc-tagged deletion constructs of G9a used in the co-immunoprecipitation assay; full length G9a protein; N-terminal of G9a protein containing the cysteine rich region; ankyrin repeats; and SET domain.

Initially, cells extracts were immunoprecipitated with MIER1 α antibody which specifically recognizes the α isoform of MIER1 and probed for G9a with 9E10 antibody. Cell extracts were also immunoprecipitated with preimmune serum, as a negative control. As evident in Fig. 7B, MIER1 α interacts with G9a in its full length form and weakly with the N-terminal; it doesn't interact with the other G9a constructs. The weak interaction seen with the N-terminal of G9a protein could be non-specific as its expression is relatively high compared to other myc-tagged G9a proteins.

Subsequently, HEK-293 cells were transiently transfected with 0.5 μ g of MIER1 α and myc-FL-G9a or myc-G9a-N-terminal or myc-G9a-ANK or myc-G9a-SET constructs; 0.5 μ g MIER1 α expressing plasmid and pCS3+MT empty vector; or 0.5 μ g myc-FL-G9a plasmid and pcDNA3.1 empty vector. This time, however, a reverse co-immunoprecipitation was performed, where G9a constructs were immunoprecipitated

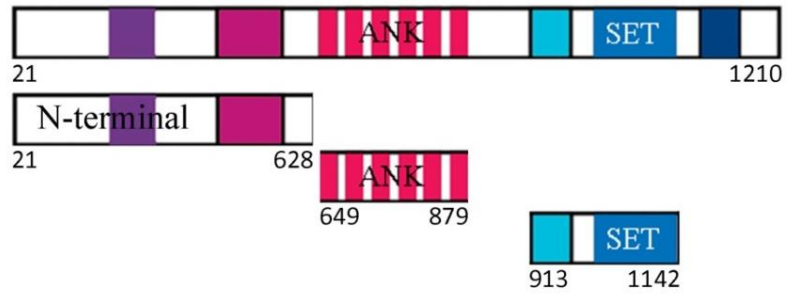
using 9E10 antibody, and probed for MIER1 α using pan-MIER1 antibody. Once again, as evident in Fig. 7C, MIER1 α only interacts with full length G9a; the deletion constructs do not interact with MIER1 α in HEK-293 cells *in vivo*.

Expression of MIER1 α and myc-tagged G9a constructs for both co-immunoprecipitations was verified by Western blot analysis of whole cell lysates (extracted in 1.5x SSB) using pan-MIER1 antibody, which recognizes both MIER1 α and MIER1 β isoforms, and 9E10 antibody which recognizes the myc-tag G9a constructs. For MIER1 α , it is the bottom band that corresponds to the protein. The predicted molecular weight of MIER1 α is 49 kDa but it runs slower on SDS-PAGE gels at 63-75 kDa. The top band which runs between 75-100 kDa, is possibly endogenous MIER1 β , which is larger than MIER1 α . MIER13A isoform is also larger than MIER1 α , which results from the inclusion of the cassette exon, exon 3A. The antibody could be detecting endogenous levels of MIER13A as the pan-MIER1 antibody used is not isoform specific. MIER1 α and MIER1 β specific antibodies could be used to determine definitely what protein the top band is. Similarly, the band can be cropped out from the gel and mass spectrometry could be performed to determine the protein. Moreover, it can also be a non-specific band that is being detected by the primary and secondary antibodies. β -actin antibody was used to probe for β -actin, which is used as a loading control for the cell lysates. The levels of β -actin are similar in all the lanes indicating that equal amount of cell lysates were loaded in each lane.

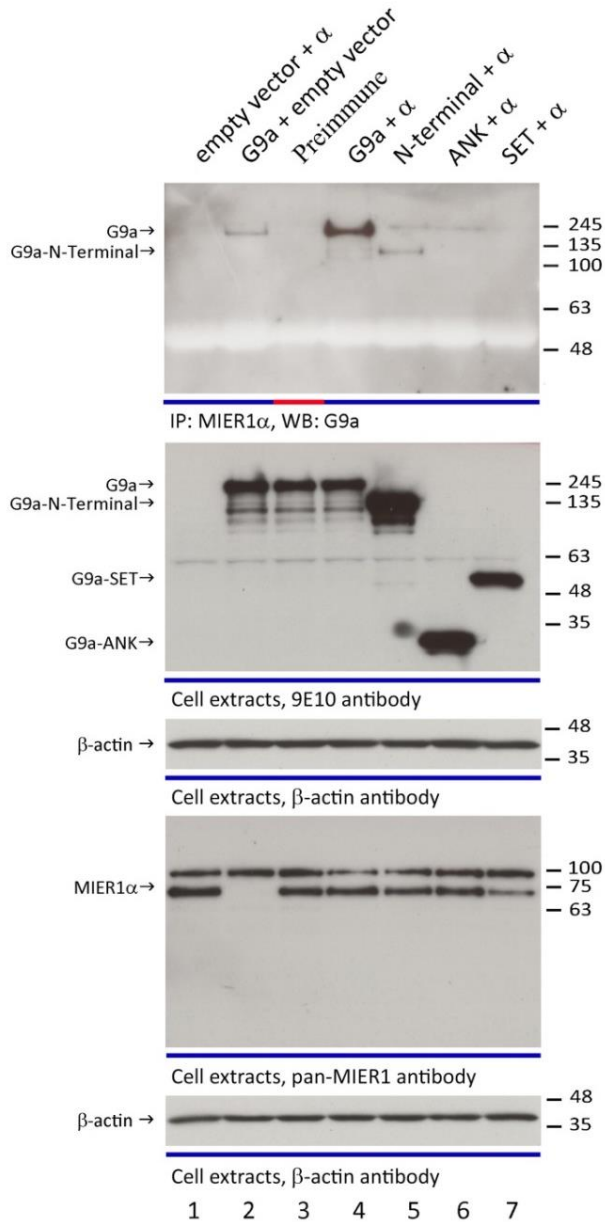
Figure 7: MIER1 α interacts with G9a in HEK 293 cells

HEK-293 cells were seeded at a density of 5×10^5 cells/well in a 6-well plate and grown in DMEM for approximately 24 hr, following which cells were transfected with 0.5 μ g of MIER1 α and myc-FL-G9a or myc-G9a-N-terminal or myc-G9a-ANK or myc-G9a-SET constructs; 0.5 μ g MIER1 α plasmid and pCS3+MT empty vector; or 0.5 μ g myc-FL-G9a plasmid and pcDNA3.1empty vector. Cell lysates were prepared as described in Materials and Methods section, and western blot was performed (lanes 1-7). **6A)** Schematic representation of the myc-tagged deletion constructs of G9a used in the assay. **6B)** The top panel shows the MIER1 α -G9a complex immunoprecipitated using MIER1 α antibody and preimmune serum (negative control), and western blotted for G9a using 9E10 antibody. The bottom western blots shows the cell extracts probed for G9a using 9E10 antibody and for MIER1 α using MIER1 antibody (the bottom band seen in the western blot is MIER1 α). They show the expression levels of the transfected proteins. β -actin antibody was used to probe for β -actin, which is used as a loading control for the cell lysates. **6C)** The MIER1-G9a complex immunoprecipitated using 9E10 antibody (immunoprecipitated myc-tagged G9a), and western blotted for MIER1 α using MIER1 antibody. The bottom western blots shows the expression level of the transfected proteins. 9E10 antibody was used to detect myc-G9a constructs and pan-MIER1 antibody to detect MIER1 α (the bottom band seen in the western blot is MIER1 α). β -actin antibody was used to probe for β -actin, which is used as a loading control for the cell lysates.

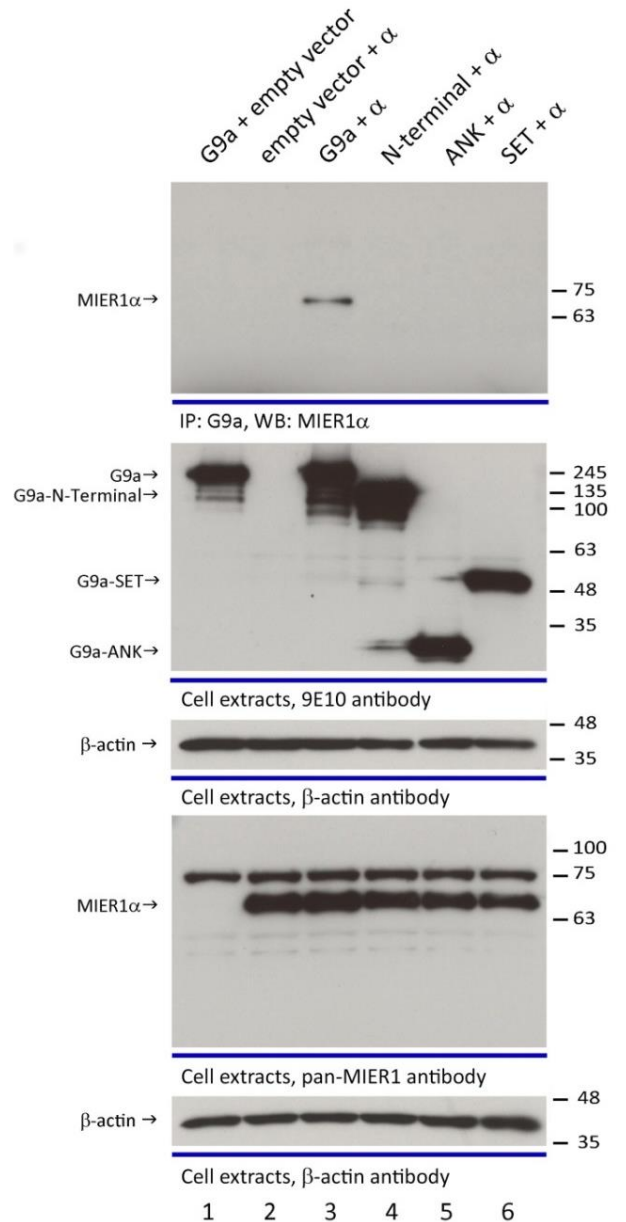
A.



B.



C.



3.2 Interaction between MIER1 α , G9a, CDYL and EZH2

3.2.1 *In vivo* interaction between MIER1 α , G9a and CDYL

Co-immunoprecipitation assay was performed between MIER1 α , G9a, and CDYL to determine whether these proteins interact and exist as a complex *in vivo*. HEK-293 cells were seeded at a density of 5×10^5 cells/well in a 6-well plate and grown in DMEM for approximately 24 hr. Cells were then transiently transfected with: 1) 0.5 μ g MIER1 α , myc-G9a and flag-CDYL plasmids; 2) 0.5 μ g MIER1 α , myc-G9a and pCMV Tag2b empty vector; 3) 0.5 μ g MIER1 α , pCS3+MT empty vector and flag-CDYL; 4) 0.5 μ g pcDNA3.1empty vector, myc-G9A and flag-CDYL; 5) 0.5 μ g MIER1 α , pCS3+MT empty vector and pCMV+tag2b empty vector. Cell lysates were immunoprecipitated using MIER1 α antibody, and a Western blot analysis was performed. The MIER1-G9a-CDYL complex was also immunoprecipitated with preimmune serum as a negative control (Fig. 8A). Additionally, the expression of the transiently transfected proteins was checked by Western blot analysis. For both, immunoprecipitation (Fig. 8A) and cell lysate western blots (Fig. 8B), pan-MIER1 antibody was used to detect MIER1 α , flag antibody was used to probe for flag-tagged CDYL, and 9E10 antibody was used to visualise myc-tagged G9a.

As evident by Fig. 8, G9a and CDYL both interact with MIER1 α . Moreover, MIER1 α interacts with CDYL independent of G9a being present or not (Fig. 8, lane 4). This is the same with MIER1 α -G9a interaction; MIER1 α interacts with G9a independent of CDYL (Fig. 8, lane 3). When CDYL is transiently transfected alongside MIER1 α and

G9a, MIER1 α -G9a interaction is neither amplified nor diminished (Fig. 8, lane 2), and vice versa, suggesting that CDYL is not a bridging protein between MIER1 α and G9a. Lanes 4 and 5 in Fig. 8 serve as additional negative controls. Lane 4 shows no G9a and CDYL binding to MIER1 α as there is no MIER1 α transfected into the cells; when the cell lysates are immunoprecipitated with MIER1 α antibody and there is no MIER1 α transfected in the cells, there should be no G9a or CDYL co-immunoprecipitated. This is evident in Fig. 8 lane 4; there is no non-specific G9a or CDYL binding to the beads. Fig. 8 lane 5 shows no G9a and CDYL binding either since neither protein is transfected; only MIER1 α is transfected. The absence of bands shows that 9E10 and flag antibody used to probe for G9a and CDYL are myc-tagged and flag-tagged specific.

The expression of MIER1a is greater in lane three compared to other lanes where it is transfected. This could be due to error in transfection of the MIER1 α -expressing plasmid. The levels of G9a expressed in each condition are similar as shown by Fig. 7B. CDYL expression, however, differs with and without the expression of MIER1 α . When MIER1 α is expressed in cells, there is increased CDYL expression. β -actin antibody was used to probe for β -actin, which is used as a loading control for the cell lysates. The levels of β -actin are similar in all the lanes indicating that equal amount of cell lysates were loaded in each lane.

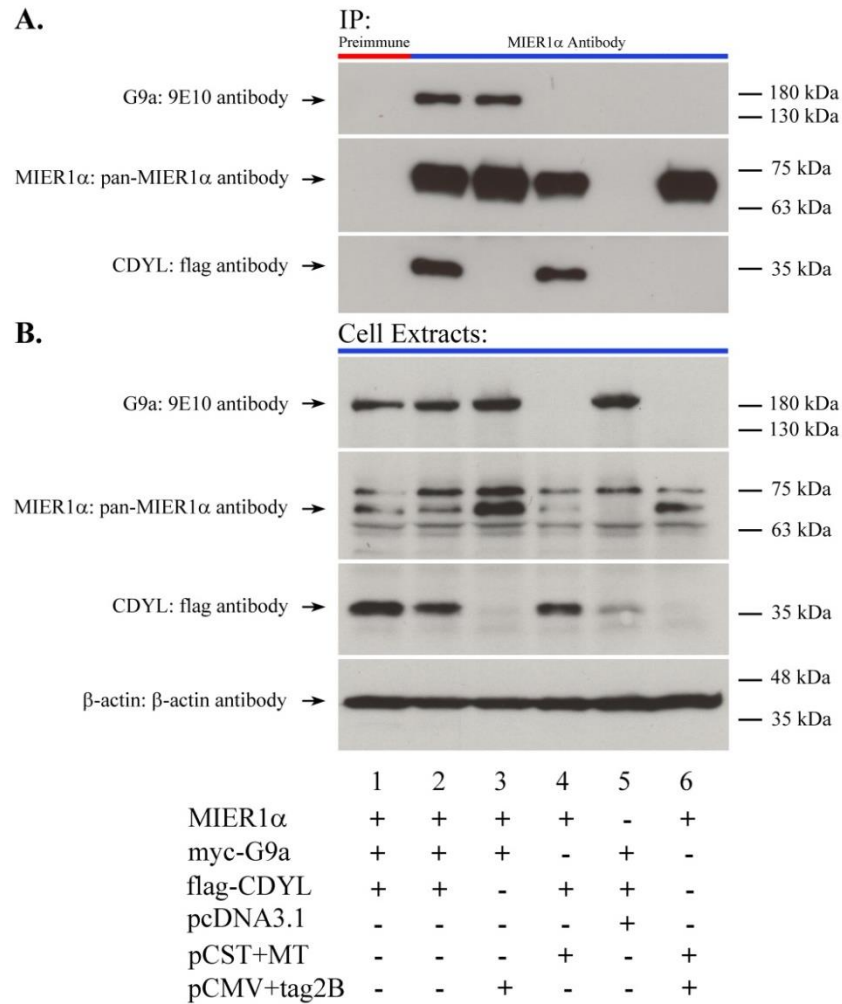


Figure 8: MIER1 α interacts with G9a and CDYL in HEK 293 cells.

HEK-293 cells were seeded at a density of 5×10^5 cells/well in a 6-well plate and grown in DMEM for approximately 24 hr. Cells were then transiently transfected with: 1) 0.5 μ g MIER1 α , myc-G9a and flag-CDYL; 2) 0.5 μ g MIER1 α , myc-G9a and pCMV-tag2b vector; 3) 0.5 μ g MIER1 α , pCST+MT vector and flag-CDYL; 4) 0.5 μ g pcDNA3.1 vector, myc-G9A and flag-CDYL; 5) 0.5 μ g MIER1 α , pCST+MT vector and pCMV+tag2b vector. Cell lysates were prepared as described in Materials and Methods section, and Western blot was performed (lanes 1-6). The co-immunoprecipitation was replicated three times. **A)** MIER1 α -G9a-CDYL complex was immunoprecipitated using MIER1 α antibody and preimmune serum (negative control). **B)** Cell lysates showing the expression level of transiently transfected proteins. Monoclonal pan-MIER1 antibody was used to detect MIER1 α , monoclonal flag antibody was used to probe for CDYL and monoclonal 9E10 antibody was used to visualise myc-G9a. β -actin antibody was used to probe for β -actin, which is used as a loading control for the cell lysates.

3.2.2 *In vivo* interaction between G9a and CDYL

Co-immunoprecipitation assay was performed between G9a and CDYL to determine whether these proteins interact *in vivo*. HEK-293 cells were seeded at a density of 5×10^5 cells/well in a 6-well plate and grown in DMEM for approximately 24 hr. Subsequently, cells were transiently transfected with: 1) 0.5 μ g myc-G9a and 1 μ g flag-CDYL plasmids; 2) 0.5 μ g pCS3+MT empty vector and 1 μ g flag-CDYL plasmid; 3) 0.5 μ g myc-G9a plasmid and 1 μ g pCMV-Tag2B empty vector. Following 48 hr, cell lysates were immunoprecipitated using 9E10 antibody which recognises myc-tagged G9a, and Western blot analysis was performed. The expression of the transiently transfected proteins was checked by Western blot analysis as well. For both, immunoprecipitation (Fig. 9A) and cell lysate western blots (Fig. 9B), flag antibody was used to probe for CDYL, and 9E10 antibody was used to visualise myc-G9a. As evident by Fig. 9A, G9a and CDYL do not interact *in vivo*. The expression of CDYL is low in comparison to G9a (Fig. 9B). However, in previous co-immunoprecipitation assay between MIER1, G9a and CDYL (Fig. 8), CDYL strongly interacted with MIER1 with low levels of CDYL protein being expressed.

Densitometry measurements were performed on the Western blot bands representing the CDYL expression levels in both figures (Fig. 8 and Fig. 9). CDYL expression in Fig. 8 is higher when MIER1 α is also expressed; with MIER1 α expression, densitometry reading of the CDYL band is 16212; without MIER1 α expression, it is 4253. CDYL expression is about 3.8 times greater when MIER1 α is also expressed in HEK-293 cells. In Fig. 9, when CDYL and G9a are expressed without MIER1 α , CDYL

expression is low with a densitometry reading of 10056, but it is higher than in Fig. 8 when only G9a and CDYL are expressed (densitometry reading of 4253). Thus, the lack of interaction seen between G9a and CDYL is not due to low levels of CDYL, since CDYL interacts with MIER1 α with similar expression levels.

β -actin antibody was used to probe for β -actin as a loading control for the cell lysates. The levels of β -actin are similar in all the lanes indicating that equal amount of cell lysates were loaded in each lane.

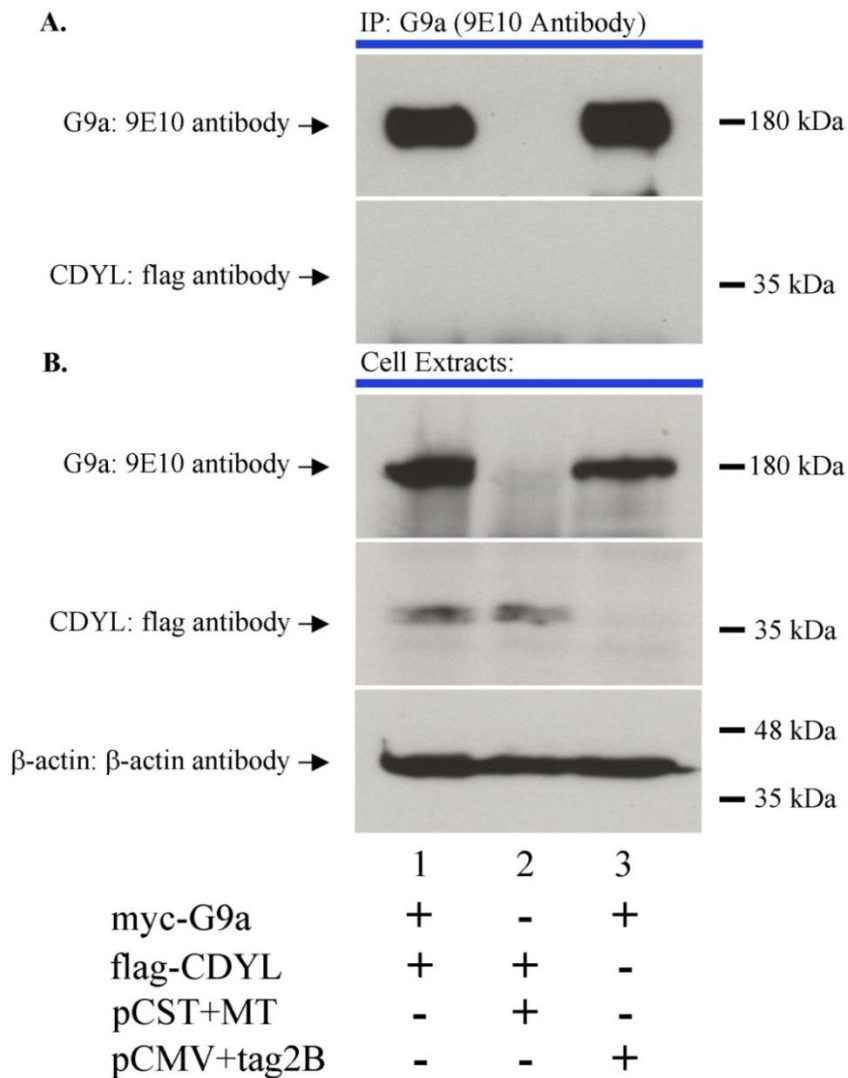


Figure 9: G9a and CDYL do not interact *in vivo* in HEK 293 cells

HEK-293 cells were seeded at a density of 5×10^5 cells/well in a 6-well plate and grown in DMEM for approximately 24 hr. Cells were then transiently transfected with: 1) 0.5 μ g myc-G9a and 1 μ g flag-CDYL plasmids; 2) 0.5 μ g pCS3+MT empty vector and 1 μ g flag-CDYL plasmid; 3) 0.5 μ g myc-G9a plasmid and 1 μ g pCMV-tag2b empty vector. Cell lysates were prepared as described in Materials and Methods section, and Western blot was performed (lanes 1-3). The co-immunoprecipitation was repeated three times. **A)** G9a-CDYL complex was immunoprecipitated using 9E10 antibody. **B)** Cell lysates showing the expression level of transiently transfected proteins. Monoclonal flag antibody was used to probe for CDYL and monoclonal 9E10 antibody was used to visualise myc-G9a. β -actin antibody was used to probe for β -actin as a loading control for the cell lysates

3.2.3 *In vivo* interaction between MIER1 α , CDYL and EZH2

Co-immunoprecipitation assay was performed between MIER1 and CDYL to determine whether CDYL acts as a bridging protein between MIER1 and EZH2 *in vivo*. CDYL has been shown to interact with EZH2 [96]. HEK-293 cells were seeded at a density of 5×10^5 cells/well in a 6-well plate. After being grown in DMEM for approximately 24 hr, the cells were transiently transfected with: 1) 0.5 μ g myc-MIER1 α and 1 μ g flag-CDYL plasmids; 2) 0.5 μ g pCS3+MT empty vector and 1 μ g flag-CDYL plasmid; 3) 0.5 μ g myc-MIER1 α plasmid and 1 μ g pCMV-tag2b empty vector. Following 48 hr, cell lysates were immunoprecipitated using 9E10 antibody which recognises myc-MIER1 α , and Western blot analysis was performed (Fig. 10A). The expression of the transiently transfected proteins was checked by Western blot analysis (Fig. 10B). Flag antibody was used to probe for CDYL, 9E10 antibody was used to detect myc-MIER1 α , and EZH2 antibody was used for EZH2 visualisation. β -actin antibody was used to probe for β -actin as a loading control for the cell lysates.

Fig. 10A shows EZH2 interaction with MIER1 α only when CDYL is transiently transfected alongside MIER1 α . This suggests that CDYL bridges EZH2, enabling it to interact with MIER1 α . Fig. 9B shows the expression level of EZH2, MIER1 α and CDYL. With the exception of CDYL, the proteins have relatively similar expression in each sample. The expression of CDYL is higher in lane 1 of Fig. 10B, than lane 2. Potential explanation for this deviation in CDYL expression could have been possible inaccuracies in transfection of the CDYL-expressing plasmid. However, this experiment was repeated three times, with equal amount of the CDYL-expressing plasmid transfected

in each condition, all of which yielded the same result. The expression of CDYL is higher when MIER1 α is transfected alongside which is also evident in Fig. 8. It is possible that MIER1 α itself could have an effect on CDYL expression. β -actin expression level is similar in all the lanes, which signifies that equal amount of cell lysates were loaded in each lane. Further experiments are needed to determine if MIER1 is responsible for this augment in CDYL expression.

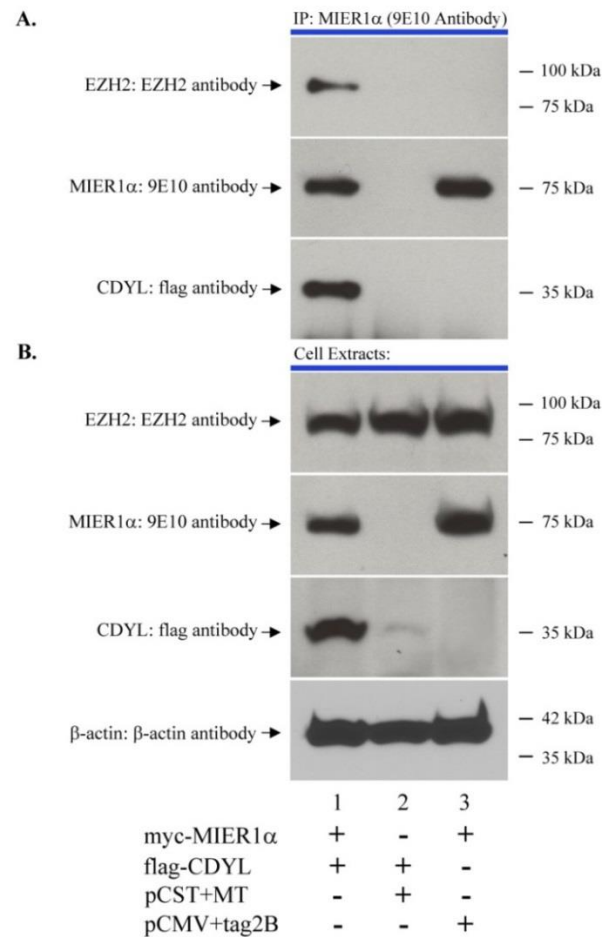


Figure 10: CDYL bridges EZH2 to MIER1 *in vivo*

HEK-293 cells were seeded at a density of 5×10^5 cells/well in a 6-well plate and grown in DMEM for approximately 24 hr. Following which, cells were transiently transfected with: 1) 0.5 μ g myc-MIER1 α and 1 μ g flag-CDYL plasmid; 2) 0.5 μ g pCST+MT empty vector and 1 μ g flag-CDYL plasmid; 3) 0.5 μ g myc-MIER1 α plasmid and 1 μ g pCMV-tag2b empty vector. Cell lysates were prepared as described in Materials and Methods section, and Western blot was performed (lanes 1-3). The co-immunoprecipitation was repeated three times. **A)** MIER1 α -CDYL complex was immunoprecipitated using 9E10 antibody. **B)** The expression level of transiently transfected proteins (MIER1 α and CDYL) and endogenous EZH2. Flag antibody was used to probe for CDYL, 9E10 antibody was used to detect MIER1 α , and EZH2 antibody was used to visualise EZH2. β -actin antibody was used to probe for β -actin as a loading control for the cell lysates.

3.3 *Histone 3 lysine 27 methylation using recombinant G9a protein*

3.3.1 Methylation assay using H3(21-44) peptide as the substrate

Recombinant G9a was used to determine whether G9a is able to methylate H3K27 via the methylation assay using the ELISA approach described in Methods and Materials section 2.6.1 (pages 62-64). The activity of G9a from two sources was tested; the SET domain from Active Motif, and the full-length G9a protein from Origene. Two concentrations of the enzyme were tested, 2.5 nM and 12.5 nM or 2.5 nM and 25 nM for the SET domain and for the full-length G9a, respectively, with 20 μ M SAM, and two different inhibitors; 200 μ M SAH and 17 μ M BIX-01294.

As evident in Figure 11, only the SET domain has some methyltransferase activity in this assay, while the full length G9a does not, suggesting that the full length recombinant protein is inactive at these concentrations. The SET domain has higher activity at the higher concentration (12.5 nM) than at the lower concentration (2.5 nM), which indicates specificity. The activity of the SET domain is attenuated in the presence of the inhibitors SAH and BIX.

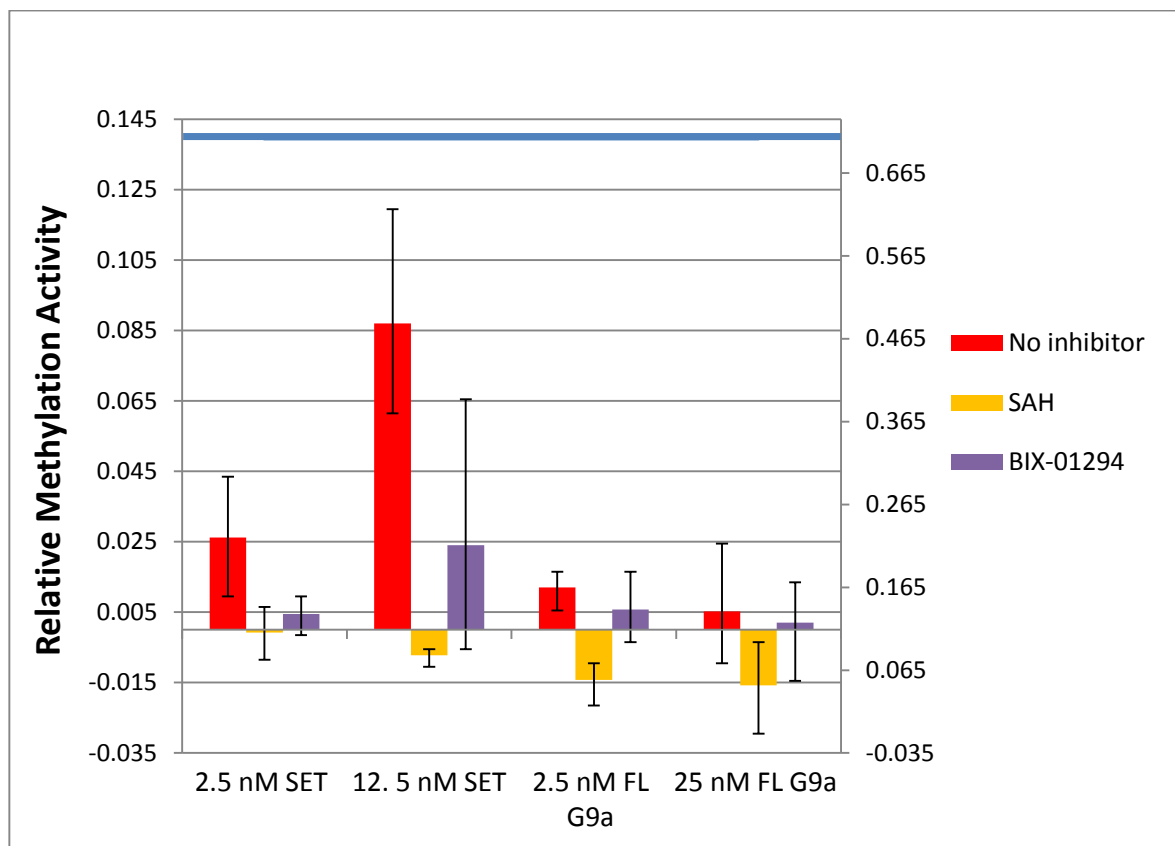


Figure 11: Recombinant G9a-SET protein weakly methylates lysine 27 on histone H3 peptide

The activity of recombinant G9a containing only the SET domain (Active motif), and the full-length protein (Origene) was tested using the Elisa method as described in Materials and Methods section. Mono- and di-methylation of the peptide substrate was detected using H3K27me1 and H3K27me2 antibodies (Active motif). Two concentration of the enzyme were tested, 2.5nM and 12.5 nM or 2.5 nM and 25 nM for the SET domain and for the full-length G9a, respectively, with 20 μ M SAM, and two different inhibitors; 200 μ M SAH and 17 μ M BIX. The methylation assay was performed with a total reaction time of 1 hr at 37 °C. Equimolar H3(21-44)K27me2 peptide (17 μ M) as the unmodified H3(21-44) peptide substrate was loaded on the streptavidin coated plates. This illustrates the maximum attainable methylation level and is represented by the blue line on the secondary y-axis. Two independent experiments were performed, each having a duplicate experiment conducted in parallel; average of the experiments is depicted in the bar graph. Error bars show the minimum and maximum values as read by the plate reader.

3.3.2 Methylation assay using recombinant histone as the substrate

The purpose of this assay was to determine whether G9a is able to mono- or dimethylate H3K27 on recombinant histone 3 (H3), with and without the methyltransferase inhibitor, SAH, which works via product inhibition. When the methyl group is transferred from SAM to the lysine on H3, SAH is produced. Thus, when SAH is initially added, it prevents the reaction from going forward and producing more SAH, following Le Chatelier's Principle [117]. H3K9 methylation was also tested as a positive control since G9a's ability to methylate H3K9 has been extensively studied [38]. The methylation assay was performed using recombinant G9a-SET protein as described in Methods and Materials section, using 44 ng G9a-SET protein, 0.75 μ g recombinant histone 3.1, 5 μ M SAM, and 150 μ M SAH.

Western blot analysis of the samples was performed to detect methylation. The unmethylated recombinant histone 3.1 used in the assay was detected using histone 3.1 antibody, as seen in Fig. 12. H3K27me1 and H3K27me2 antibodies were used to detect H3K27 methylation, and H3K9me2 antibody was used to check for H3K9 methylation. G9a-SET protein methylates H3K9 as seen in Fig. 12. This methylation is attenuated in the presence of SAH. G9a-SET protein is not able to methylate H3K27 as there is no band detected with H3K27me1 and H3K27me2 antibodies. H3 substrate is present but there is no mono- and di-methylated H3K27 detected.

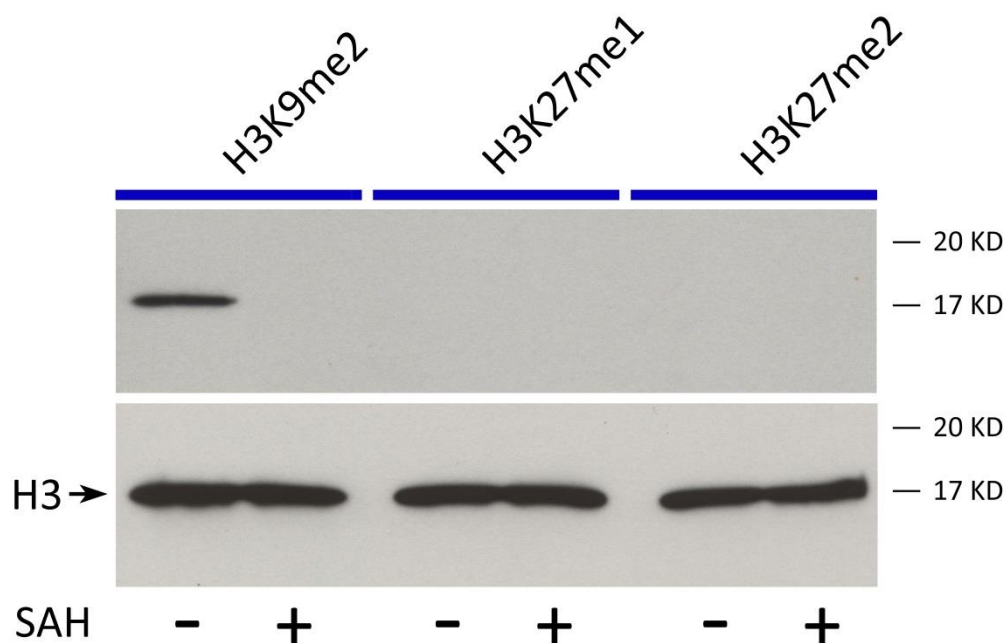


Figure 12: Recombinant G9a-SET protein methylates lysine 9 but not 27 on histone H3

The activity of recombinant G9a containing only the SET domain (Active motif) on H3K9 and H3K27 was tested and detected via western blot analysis, as described in Materials and Methods section. This assay was performed using 44 ng SET domain, 0.75 μ g recombinant histone 3.1, 5 μ M SAM, and 100 μ M SAH, in methylation assay buffer (50 mM Tris-HCl pH 9, 1 mM DTT, 0.01 % BSA, 1 mM PMSF) with a total reaction volume of 30 μ l. The total reaction time was 2 hr at 37 °C. Histone H3 antibody was used to detect H3 used in the assay. H3K27me1 and H3K27me2 antibodies were used to detect H3K27 mono- and di-methylation, and H3K9me2 antibody was used to probe for H3K9 di-methylation. H3K9 methylation was tested as a positive control.

3.3.3 Methylation assay using recombinant histone 3 (H3) as the substrate: effects of different inhibitors on G9a activity

Recombinant G9a-SET protein was used to test the potency of different inhibitors. H3K9 methylation was solely tested as previous experiment showed no H3K27 methylation by SET-G9a. The assay was performed as described in Methods and Materials section, using 44 ng G9a-SET protein, 0.75 µg recombinant histone 3.1, 5 µM SAM, and three difference inhibitors, 150 µM SAH, 170 µM BIX-01924, and 40 nM GSK-343. SAH is a common methyltransferase inhibitor (acts via product inhibition), BIX-01294 is a G9a specific inhibitor, while GSK-343 is EZH2 methyltransferase specific inhibitor.

Western blot analysis of the samples was performed and H3K9me2 antibody (1 in 1000 dilution of 1 µg/µl stock, Millipore, 07-441) was used to detect H3K9 di-methylation. A 10 % input of the recombinant histone 3.1 used in the assay was also loaded, as seen in Fig. 13, lane 1, which was probed for using histone 3.1 antibody. Moreover, G9a di-methylates H3K9 as evident by the presence of band detected by H3K9me2 antibody. SAH and BIX-01294 inhibit G9a activity and thus no band is detected on the western blot. GSK-343 is not a G9a inhibitor, as it is a specific EZH2 methyltransferase inhibitor. Thus, in the presence of GSK-343, G9a activity is not diminished and H3K9me2 is still detected.

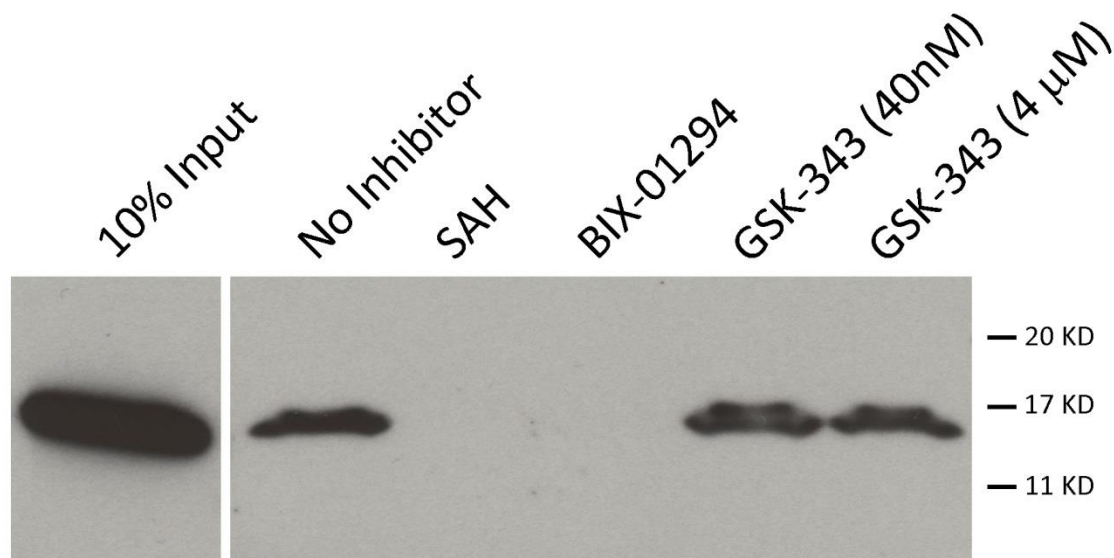


Figure 13: The activity of recombinant G9a-SET protein is inhibited by S-adenosyl homocysteine (SAH), and the G9a specific inhibitor, BIX-01294, but not GSK-343.

The activity of recombinant G9a-SET protein (Active motif) on H3K9 with and without inhibitors was tested and detected via western blot analysis, as described in Materials and Methods section. The assay was performed using 44 ng SET domain, 0.75 μ g recombinant histone 3.1, 5 μ M SAM, 100 μ M general inhibitor SAH, 170 μ M G9a specific inhibitor BIX-01294, and 40 nM and 4 μ M EZH2 specific inhibitor GSK-343, in methylation assay buffer (50 mM Tris-HCl pH 9, 1 mM DTT, 0.01 % BSA, 1 mM PMSF) with a total reaction volume of 30 μ l. Total reaction time was 2 hr at 37 °C. Histone H3 antibody was used to detect H3 used in the assay (10 % input), and rabbit polyclonal H3K9me2 antibody was used to probe for H3K9 methylation.

3.4 Histone 3 lysine 27 methylation using GST-MIER1 β

To determine whether MIER1 itself has innate methyltransferase capabilities, GST-tagged MIER1 β was used to check H3K27 methylation using ELISA. GST and GST-MIER1 β were prepared and the assay was performed as described in the “Methylation Assay using GST-MIER1 β Protein” section in Methods and Materials chapter. Two independent experiments were carried out, each having a duplicate experiment performed in parallel. The results were normalized to sample containing only the substrate peptide and graphed as histogram, Fig. 14.

Equimolar H3(21-44) K27me2 peptide (17 μ M) as H3(21-44) peptide lacking any other PTMs was loaded on the streptavidin coated plates, which represents the maximum attainable amount of methylation on the histone peptide. In other words, if all the H3(21-44) peptide used in the assay is methylated by GST-MIER1 β , it will reach the level indicated by the blue line going across the graph in Fig. 14. There is no significant difference between GST protein by itself (negative control) and GST-MIER1 protein. This indicates that MIER1 β does not methylate H3K27. However, there is no positive control in this assay to show that the assay itself was functional.

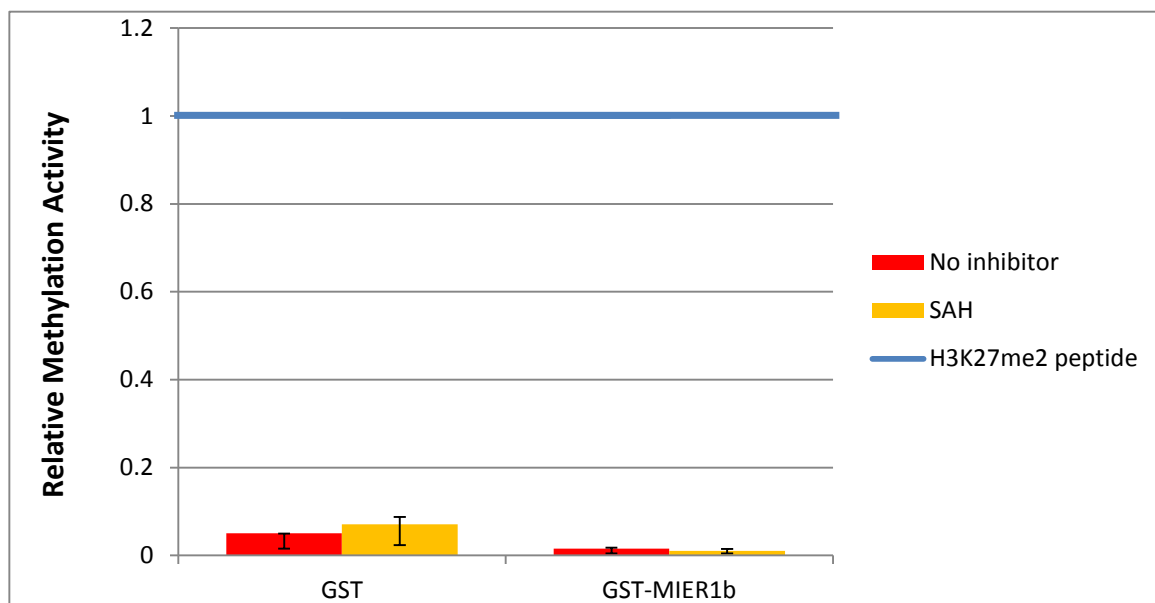


Figure 14: GST-MIER1 β does not methylate lysine 27 on histone H3 peptide

The activity of GST-fused MIER1 β was tested using the ELISA method as described in Materials and Methods section, and compared to the GST protein by itself. GST and GST-MIER1 were tested under two conditions; with and without inhibitor; 20 μ M SAM, and 200 μ M SAH. The methylation assay was performed with a total reaction time of 1 hr at 37°C. Mono- and di-methylation of the peptide substrate was detected using H3K27me1 and H3K27me2 antibodies (Active motif). Equimolar H3(21-44) K27me2 peptide (17 μ M) as H3(21-44) peptide substrate was loaded on the streptavidin coated plates; illustrates the maximum attainable methylation activity (blue line). Two independent experiments were performed, each having a duplicate (performed in parallel) within experiment; average of the experiments is depicted in the bar graph. Error bars show the minimum and maximum values as read by the plate reader.

3.5 Histone 3 lysine 27 methylation by endogenous MIER1 α complex

HEK-293 cells were transfected with empty myc (pCS3+MT) vector or myc-MIER1 α plasmid. Initially, cell lysates were collected and the MIER1 α complex was immunoprecipitated using MIER1 α antibody. Subsequently, the immunoprecipitated MIER1 α complex was used in methylation assay described in Materials and Methods section. The assay was normalised with sample containing only the H3(21-44) peptide (the amount of methylation activity detected by the plate reader in the sample containing just the substrate [H3(21-44) unmodified peptide was subtracted from each sample] and the results were graphed as shown in Figure 15. The purpose of this assay was to check if the MIER1 α complex immunoprecipitated from HEK-293 cells can methylate H3K27.

The negative control shows the methylation activity of samples containing immunoprecipitated lysates from cells transfected with empty myc (pCS3+MT) vector. Equimolar H3(21-44) K27me2 peptide (17 μ M) as H3(21-44) peptide lacking any PTMs was loaded on the streptavidin coated plates, representing the maximum attainable amount of methylation on the histone peptide. The blue horizontal line in Fig. 15 indicates the methylation level achieved if all of the H3(21-44) peptide substrate used in the assay is methylated.

The cell lysates were initially immunoprecipitated with MIER1 α antibody or with preimmune serum. The preimmune serum acts as an additional negative control. The methylation assay was performed with cell lysates containing empty vector or MIER1 α , with or without inhibitors (SAH or BIX-01294).

As evident in Fig. 15, there was no significant difference between the negative control and other conditions when immunoprecipitated with the preimmune serum. However, there is significant difference between the negative control and other conditions when the cell lysates are immunoprecipitated with MIER1 α antibody. There is some activity seen in the negative control, which has only empty vector transfected, when immunoprecipitated with MIER1 α antibody. This could be due to endogenous MIER1 α being immunoprecipitated and thus showing methylation activity. However, there is no significant difference between the negative control group when immunoprecipitated with preimmune or with MIER1 α antibody. Furthermore, there is no difference between the methylation activity of MIER1 α without inhibitors or with SAH or BIX-1294. When MIER1 α complex is immunoprecipitated from HEK-293 cells, and subsequently used in methylation assay, it has methylation activity towards H3K27. However, the inhibitors are not able to inhibit the methyltransferase activity at the concentration used; 200 μ M SAH, and 17 μ M BIX-01294. Higher concentration of inhibitors might be needed to inhibit the methylation seen. Optimization of this assay is further needed to determine the role of MIER1 complex on H3K27 methylation with certainty.

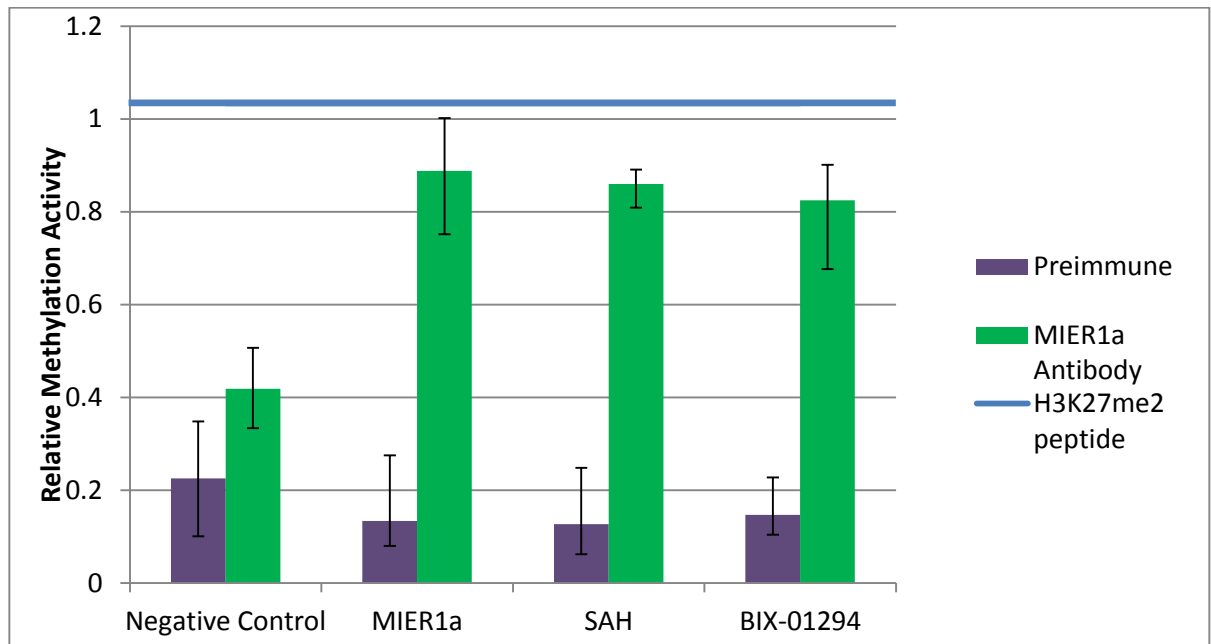


Figure 15: MIER1 α complex methylates lysine 27 on H3(21-44) peptide

Cells were seeded at a density of 5×10^5 cells /well in 6 well plates and grown in DMEM for 24 hr. Cells were then transfected with 0.5 μ g pCS3+MT empty vector or myc-MIER1 α plasmid. Cell extracts were initially immunoprecipitated with either preimmune serum or MIER1 α antibody and then used in the methylation assay as described in Materials and Methods section. The assay was performed using 5 μ M SAM, with and without the inhibitors, 200 μ M SAH and 17 μ M BIX-01294, with a total reaction time of 1 hr at 37°C. Equimolar H3(21-44) K27me2 peptide (17 μ M) as the unmethylated H3(21-44) peptide was loaded on the streptavidin coated plates which represents the maximum amount of methylation that could occur on H2(21-44) peptide. This assay was performed in duplicates twice; average of the experiments is depicted in the bar graph.

Chapter 4: Discussion

4.1 General discussion

The purpose of this study was to investigate the role of MIER1 and G9a on histone methylation of lysine 27 on histone H3 (H3K27). Recently, MIER1 has been found to recognise and specially bind to acetylated H3K27 residue (Paterno, unpublished data). Our lab has also shown MIER1 to target HDAC's activity towards H3K27ac. The presence of MIER1 augments HDAC activity to remove the acetyl group from H3K27ac (Paterno, unpublished data). Since MIER1 has a role in H3K27 deacetylation, in this study we investigated the impact of MIER1 on H3K27 methylation.

In 2008, Mulligan *et al.* [93] carried out an affinity purification of Flag-HA-tagged CDYL from HeLa nuclear extracts to isolate CDYL and its associated proteins, in an attempt to determine the mechanism and biological function of CDYL. They performed a mass spectrometry analysis and identified more than 22 associated proteins that copurified with CDYL. These included: HDAC1, HDAC2, G9a, EuHMT1 (G9a-like protein, GLP), WIZ (widely interspaced zinc fingers), REST (RE1-binding silencer of transcription) and MIER1. Subsequently, they performed a glycerol gradient sedimentation analysis to determine whether these proteins form single or discrete subcomplexes. From the results, the authors concluded that there are at least two multiprotein subcomplexes. One subcomplex contains MIER1/2, HDAC1, and HDAC2, while the second slower sedimenting complex includes REST, WIZ, G9a, and GLP. They also concluded that the results suggest CDYL and REST and histone methyltransferases

to be in the same complex. However, their results show CDYL to be present in both sedimentary fractions, with higher amounts in the fraction including the first multiprotein complex. Similarly, even though the authors conclude that G9a belongs to the second complex, it does overlap with the first complex. Furthermore, MIER1 follows the same trend as CDYL, and is present in both multiprotein complexes [93].

The atrophin family proteins including arginine-glutamic acid dipeptide repeats protein (RERE) and *Drosophila* Atrophin (Atro) are nuclear receptor corepressors which share the ELM2 and SANT domains with MIER1 [116]. In 2008, Wang *et al.* demonstrated that both RERE and Atro interact directly with G9a via the SANT domain. The authors showed MIER1 to interact with G9a as well [116]. Wang *et al.*, however, did not demonstrate which region of MIER1 is responsible for the MIER1-G9a interaction. Thus, MIER1-G9a binding was further investigated in my study.

Initially, a GST pull-down assay using GST-MIER1 constructs and recombinant G9a was performed to determine whether MIER1 and G9a interact *in vitro*. The GST-MIER1 constructs used were full length MIER1 β , MIER1 β with a three point mutation (F326A, Y327A, Y328A) in the SANT domain, and three deletion constructs; GST-MIER1 β construct aa1-283, which contains the ELM2 domain, but lacks the SANT domain; GST-MIER1 β construct aa287-512, which contains the SANT domain, but lacks the ELM2 domain; and GST-MIER1 β construct aa287-357, which contains the SANT domain with an additional 25 aa downstream of it. GST was used as a negative control.

The amino acids 326-328 were mutated as they are highly conserved in evolution and mutagenesis of these aa leads to changes in the SANT structure [118]. Alignment of

MIER1 SANT domain from 9 species, mouse, rat, human, cow, chicken, *X. laevis*, *X. tropicalis*, zebrafish, and rainbow trout, showed the SANT domain to display 88 % identity, with the three aa 326-328 (FYY) to be among the 100 % conserved residues [118]. Moreover, the SANT domain contains three α -helices which contain highly conserved bulky aromatic residues. These aromatic residues form a hydrophobic core essential for the overall structural fold of the SANT domain (reviewed in [119]). Secondary structure predictions of the MIER1 SANT domain show aa 326-328 to be included in the third α -helix. Thus, mutagenesis of these aa to alanine would disrupt the hydrophobic core of the SANT domain and lead to overall structural alterations.

The results (Fig. 6) show MIER1 β to interact with G9a, and this interaction is only evident with the GST-MIER1 β constructs containing the SANT domain. GST-MIER1 β construct with the 3 point mutation in the SANT domain interacts with G9a as well, suggesting that these amino acid residues do not play a role the MIER1-G9a interaction. Moreover, GST by itself interacts with G9a; however, relatively greater amounts of GST protein was utilized in the assay. The relatively high amount of GST protein could be causing G9a to bind non-specifically. Densitometry measurements of the blots were performed for both the amount of G9a binding and the levels of GST-fusion constructs used. The amount of G9a binding was then divided by the level of GST proteins used and the results were graphed. As evident by Fig. 6B, MIER1 β interacts with G9a directly. G9a interaction with GST protein is low in comparison to MIER1 β (Fig. 6C).

Following the *in vitro* GST pull down, co-immunoprecipitated between MIER1 α and G9a was performed to elucidate if the two proteins interact and exist in a complex *in vivo*. Initially, HEK-293 cells were transiently transfected with MIER1 α and G9a, and MIER1 α was immunoprecipitated from the cells using MIER1 α antibody. The results show that MIER1 α interacts with full length G9a protein; deletion construct of the protein (N-terminal, Ankyrin repeats, SET domain) do not bind to MIER1 α (Fig. 7B). Subsequently, a reverse co-immunoprecipitation was performed where myc-G9a was immunoprecipitated from the cells using 9E10 antibody, and whether MIER1 α co-purifies with G9a was tested. Once again, the results show MIER1 α interacts with full length G9a protein only (Fig. 7C) [116].

The GST pull down assay and the co-immunoprecipitation assays show that MIER1 interacts with G9a, both *in vitro* (Fig. 6) and *in vivo* (Fig. 7). As evident by the results, the interaction is direct and it is the region of MIER1 β containing the SANT domain that is responsible for the binding between the two proteins. GST pull-down between G9a and GST-fused MIER1 β construct lacking only the SANT domain is required to conclude with certainty that MIER1 β interacts with G9a via this domain. This is consistent with other transcriptional repressors, such as Atro and RERE, which also bind G9a by their SANT domain.

According to Mulligan *et al.* (2008) [93], G9a, REST and CDYL exist in one complex, while MIER1 exists in a complex with HDAC1 and HDAC2. However, as shown by our results, MIER1 α interacts with G9a, both *in vitro* and *in vivo*. Similarly, we tested the possibility of CDYL having a role in MIER1 α -G9a interaction (Fig. 8). Co-

immunoprecipitation between MIER1 α , G9a and CDYL (Fig. 8) showed that the presence of CDYL did not augment or attenuate the level of G9a interacting with MIER1 α . Similarly, the presence of G9a did not affect the level of CDYL interacting with MIER1 α . This suggests that MIER1 α interacts with G9a and CDYL independently of each other.

To further validate the point that CDYL does not have a role in MIER1-G9a interaction, CDYL-G9a interaction was tested. Mulligan *et al.* [93] had tested CDYL, REST, and G9a interaction by performing a co-immunoprecipitation. They immunoprecipitated REST and checked for the presence of CDYL and G9a. Their results indicated that CDYL bridges G9a to REST. They concluded that REST, CDYL and G9a are components of the same protein complex [93]. Additionally, they performed a GST pull down assay between GST-fused CDYL and flag-tagged G9a and their results showed G9a to directly interact with GST-fused CDYL, *in vitro*. They did not perform any *in vivo* assays to verify this interaction. However, when we immunoprecipitated myc-tagged G9a from HEK293 cells, CDYL did not co-purify with it, suggesting that they do not interact *in vivo* (Fig. 9). The levels of flag-tagged CDYL expressed in the cells was low, however, it is relatively equal to the co-immunoprecipitated assay between MIER1 α , G9a and CDYL (Fig. 8). In that assay, CDYL interacted with MIER1 α strongly despite the low CDYL expression level. Mulligan *et al.* (2008) [93] did not check for the presence of MIER1 in the co-immunoprecipitation performed between REST, CDYL, and G9a. As evident by the western blot analysis of the glycerol gradient gel assay performed by the authors, MIER1 is present in both of the fraction peaks of the two multiprotein

subcomplexes. Thus, in their immunoprecipitation assay, it is possible that endogenous MIER1 was bridging G9a to CDYL, which in turn bridged G9a to REST.

In 2011, Zhang *et al.* [96] showed that CDYL recognizes di- and tri-methylated H3K27 and directly interacts with EZH2, which is the methyltransferase component of PRC2 complex. Initially, they performed histone binding assays using native calf thymus total histones (CTH) as substrates, which demonstrated that recombinant CDYL protein mainly interacted with histone H3, but not H4, H2A, or H2B. To determine whether PTMs of H3 have an impact on CDYL-H3 interaction, the authors compared the binding affinity of CDYL to native CTHs, which have multiple *in vivo* modifications, and to recombinant *Xenopus* histone octomers, which lack any PTMs. Their results showed CDYL to only interact with native H3 and not with the recombinant H3 octomers. This suggested that PTMs of H3 play an important role in the recruitment of CDYL. To further verify this, and to determine which PTMs are crucial for this interaction, they performed histone-peptide binding assays. Their results showed CDYL to strongly interact with H3K9me3, H3K27me2, and H3K27me3 only [96]. H3K27 di- and tri-methylation marks are catalyzed by the PRC2 complex member, EZH2. Thus, Zhang *et al.* (2011) [96] examined the *in vivo* interaction between CDYL and PRC2 by immunoprecipitating CDYL from MCF7 cells. Western blot analysis was performed on the immunoprecipitates using antibodies against the core components of PRC2. The authors showed that EZH2, SUZ12 and EED, which comprise the PRC2 complex, all interact with CDYL *in vivo*. Following the immunoprecipitation assays, the authors performed GST pull-down assays between CDYL and the components of PRC2 complex. The GST pull-down assays

showed CDYL to directly interact with EZH2 but not EED and SUV12 [96]. They concluded that PRC2 is recruited by CDYL *in vivo* by through interaction with EZH2.

Recently, two studies have reported EED, a component of PRC2, to bind to histone tails with Kme3 PTM by its carboxyl-terminal WD40, and to be responsible for H3K27me3 propagation on neighboring histones. Margueron *et al.* (2009) [120] and Xu *et al.* (2010) [121] showed that EED binds to Kme3 on histone peptides via the top face of its β -propeller architecture. Three aromatic residues form an aromatic cage which accommodates the ammonium group of the Kme3, while a fourth aromatic residue flanks the aliphatic chain of Kme3. Their structural data elucidates the preferential recognition EED has for the the Ala-Arg-Lys-Ser motif-containing H3K27me3, H3K9me3, and H1K26me3 marks over lower methylation states and other histone methyl lysine marks [120, 121]. Margueron *et al.* (2009) [120] also showed that EED binding to histone tails with K27me3 marks leads to the allosteric activation of the methyltransferase activity of PRC2 complex. Initially, they performed histone lysine methyltransferase (HKMT) assays using recombinant oligonucleosomes using unmodified or methylated H3K27 peptides. Their results showed that using unmodified or mono-methylated H3K27 peptides as substrates did not affect the enzymatic activity of PRC2 significantly, but tri-methylated peptides increased its activity by 7 fold. To ascertain whether the observed stimulation by the tri-methylated peptides was EED-mediated, the authors used PRC2 complex containing EED with mutations in the residues forming the aromatic cage (Phe-97Ala or Tyr-365Ala). These mutations prevent EED from recognising and binding to the repressive tri-methylated lysines. When they used the PRC2 complex with mutant EED to

perform the HKMT assays, the basal activity was similar to the wild-type, but the stimulated activity seen with the addition of H3K27me3 peptides was abolished. Moreover, in *Drosophila*, mutated EED caused a reduction in global methylation and disrupts development [120]. Xu *et al.* (2010) [121] also showed that EED binds to its ligands weakly (K_d : 108 μ M with H3K27me3) compared to the HP1 chromodomain (K_d : 4 μ M with H3K9me3) and the JMJD2A tudor domains (K_d : 0.4 μ M with H3K4me3), which suggests that additional chromatin readers might play a role in the recruitment of PRC2 complex. The binding affinity of CDYL for H3K27me3 is approximately 20-fold higher than that of EED [96]. In addition to binding to H3K27me3, CDYL also recognises and binds H3K27me2, while EED does not [96]. Together, these studies indicate an additional putative chromatin reader, which connects the unmodified H3K27 residue to CDYL and subsequently to the PRC2 complex.

Paterno *et al.* have shown MIER1 β to specifically recognize and bind to H3K27ac (unpublished data). MIER1 β has also been shown to recruit HDAC [106] and augment its deacetylase activity, removing the acetyl group from H3K27 (unpublished data). H3K27 can be methylated only when it is not acetylated. Thus, MIER1 presents itself as the initial chromatin reader, which recruits HDAC to remove the acetyl group and enables H3K27 to be methylated by methyltransferases such as EZH2. In previous studies [93, 122] and ours, MIER1 α interacts with CDYL. Additionally, CDYL has been shown to interact with EZH2 with higher affinity than EED [96]. Thus, we performed co-immunoprecipitation using MIER1 α antibody and checked whether EZH2 co-purifies with MIER1 α in the presence or absence of CDYL. This study showed EZH2 to interact

with MIER1 α only when there is CDYL transfected in the HEK-293 cells alongside MIER1 α (Fig. 10). This suggests that CDYL acts as a bridging protein between MIER1 α and EZH2.

EZH2 plays an important role in H3K27 tri-methylation [98, 123, 124]. However, the mechanism behind initiation of H3K27 methylation (H3K27 mono- and di-methylation) is still unclear. We hypothesized G9a to be the enzyme responsible for mono- and di-methylation of H3K27. As previously stated, MIER1 specifically recognizes and binds to H3K27ac (Paterno, unpublished data). The MIER1 recruits HDAC to H3K27ac, augmenting HDAC's activity towards H3K27ac. The initial hypothesis was that once HDAC has removed the acetyl group off H3K27, G9a subsequently mono- and di-methylates the lysine residue.

Initially, we used biotinylated H3(22-44) peptide which contains the lysine 27 residue as the substrate for the *in vitro* HMT assays. We used streptavidin-coated plates to then capture the biotinylated peptide and used antibodies against H3K27me1 and H3K27me2, followed by ELISA substrate, to detect methylation. Our initial experiment which used the ELISA approach showed G9a-SET domain to weakly methylate H3K27 (Fig. 11). However, when we used recombinant histone H3 mononucleosome without any PTMs and checked for methylation via western blot analysis, G9a was unable to methylate H3K27 (Fig.12). It did however methylate H3K9, a substrate G9a is known to methylate with high efficiency. Other studies have shown G9a to weakly mono- and di-methylate H3K27 [38, 125], but we weren't able to achieve similar results.

In 2001, Tachibana *et al.* [38] tested the HMTase activity of G9a. They demonstrated that G9a is able to methylate H3 at K9 and K27, along with H1 and other non-histone proteins. Initially, they used native and recombinant histones to investigate if they were a substrate for G9a methylation. Once established that G9a had higher activity towards histone H3, they subsequently used GST fusion histone mouse H3 amino terminus (residues 1-57) as the substrate for *in vitro* methylation assays to investigate which lysine residue G9a specifically methylates. Their results showed G9a to methylate H3K9 and H3K27, with H3K9 being the preferred target site [38]. This is consistent with the first methylation assay we performed, using biotinylated H3 peptide containing the amino terminus (residues 21-44). SET domain of G9a was weakly able to methylate H3K27 on the peptide substrate (Fig. 11). It is only when recombinant histone 3 was used as substrate that G9a was unable to methylate K27 (Fig. 12). A recent study has shown wild type G9a to mono-methylate H3K27 *in vitro*. However, in their *in vitro* HMTase assay they used reconstituted nucleosomes with recombinant histones, free of PTMs [125]. These studies suggest that G9a could mono-methylate H3K27, but our data does not follow this trend. It is possible that G9a needs the histone substrate to be more accessible or be similar to *in vivo* conditions such as being part of a nucleosome.

Nuclear receptor SET domain-containing (NSD2) histone methyltransferase has different target preferences depending on the nature of the substrate provided. When nucleosomes are the substrate, NSD2 mono- and di-methylates H3K36. However, when octomers are used as substrate, NSD2 mono- and di-methylates H4K44 [126]. This discrepancy in target preference with octomers is diminished when short double-stranded or single-stranded DNA is added, which stimulates the NSD2 activity on H3 and

inhibitors its activity on H4 [126]. Overall, DNA acts as an allosteric effector of NSD2 which makes H3K36 to become the preferred target. It is possible that G9a target preference varies with the type of substrate or presence of DNA as well, which could explain the discrepancy in H3K27 methylation.

A study done by Rea *et al.* (2000) [27] demonstrated that the SET-associated cysteine rich regions play an important role in methylation as the mutant construct of SUV39H1 lacking these regions lacks HMTase activity. Furthermore, the authors generated GST fusion products of extended SET domains of human EZH2 methyltransferase and human trithorax homolog HRX methyltransferase that correspond to GST-SUV1(82–412). Neither EZH2 nor HRX had any detectable HMTase activity towards free histones. They also checked the HMTase activity of GST-Suv2(157-477), and it was as active as GST-SUV1(82-412) [27]. The C-terminal cysteines are absent in EZH2. Similarly, the HRX lacks the SET domain-associated, cysteine-rich region. Suv39H1 and Suv39H2 have both of these cysteine domains. Together their results suggest that HMTase activity towards free histones requires both the SET domain and its adjacent cysteine-rich regions. The G9a protein we used in our *in vitro* assays is comprised of the pre- and post-SET domains containing the adjacent cysteine residues and the SET-domain itself. Thus, the enzyme is functional and its activity is evident by its ability to methylate H3K9. Nonetheless, G9a is unable to methylate H3K27 in our *in vitro* HMTase assay.

It is possible that G9a requires the presence of additional proteins, such as GLP, its dimer partner, in order to methylate H3K27. When the methyltransferase activity of G9a or GLP is attenuated, mono- and di-methylation of H3K9 is greatly reduced [45].

This might apply to H3K27 methylation as well. Furthermore, G9a and GLP are highly similar in structure, containing the same domains [46]. MIER1 interacts with G9a (Fig. 6) and given the similarities between G9a and GLP, MIER1 could also interact with GLP. Moreover, there have been studies which have demonstrated that certain proteins are functional or their activity increases significantly when in a complex [98, 124]. For example, purified flag- tagged EZH2 protein methylates histone H3, but its activity is at least a thousand fold lower than that of the EZH2-EED complex on a chromatin substrate [124]. Moreover, purified, flag-tagged EZH2 has four times greater activity on free histones than on chromatin. Overall, the difference in specific activity between purified flag-tagged EZH2 protein and the EZH2-EED complex on a histone substrate is approximately 60-fold [124]. Furthermore, recombinant EZH2 made in *Escherichia coli* or baculovirus-infected SF9 cells lacks HMTase activity [98]. This suggested that other members of the complex are needed for the HMTase activity of EZH2 [98], which could be the case with G9a.

This study demonstrated G9a to be in a complex with MIER1 α and CDYL, which is consistent with other studies [122]. A chemoproteomic profiling study of HDAC performed by Bantscheff *et al.* (2011) [122], using a combination of affinity capture and quantitative mass spectrometry, also revealed G9a, GLP, WIZ, CDYL, HDAC1 and MIER1 to exist in a complex [122]. Thus, the next step in our study was to examine the methylation activity of MIER1 towards H3K27.

Generally, HKMTs have a catalytic SET-domain which is responsible for adding the methyl group on the lysine residues. However, there is a second class of HKMTs that

do not possess a SET domain, which is currently only comprised of the evolutionary conserved protein DOT1 (disruptor of telomeric silencing, also called kmt4) [127]. MIER1 lacks the SET-domain that is present in G9a and other HKMTs, and in this study we demonstrated that MIER1 β lacks methyltransferase activity (Fig. 14). However, the experiment lacked a positive control and thus we cannot conclude with certainty that MIER1 β lacks methyltransferase activity. There is a possibility that the methylation assay used was not functional. Nonetheless, this study showed the MIER1 α complex to have methylation activity towards H3K27 (Fig.15). The assay still needs optimization as the methyltransferase inhibitor, SAH, which acts by product inhibition, was ineffective. Maybe a higher concentration of SAH is required to inhibit the reaction. Development of another methylation assay using recombinant histone as the substrate could be used in future to confirm the role of MIER1 α complex in H3K27 methylation. Overall, our results show G9a to have little to no methyltransferase activity towards H3K27. MIER1 β also does not methylate this residue on histone 3.

4.2 Proposed role of MIER1 in H3K27 methylation

MIER1 recognizes and binds to H3K27 only when the lysine is acetylated (Paterno, unpublished data). Additionally, MIER1 exists in a complex with HDAC1 [106]. Thus, MIER1 recruits this histone-modifying enzyme, and enables it to remove the acetyl group from lysine 27 on H3 (Fig. 16A). Recent findings have shown the presence of MIER1 to increase HDAC1 activity significantly (Paterno, unpublished data). By removing the acetyl group from the lysine residue, the lysine is now available to be methylated by HKMTs. Initially, the hypothesis was that MIER1 would then recruit G9a,

which would mono- and di-methylate H3K27. However, this study showed G9a to have no detectable methyltransferase activity towards H3K27.

It is possible that EZH2 is responsible for H3K27 methylation or that G9a needs its binding partners to function. In 2006, Schoeftner *et al.* [128] investigated the role of PRC2 in H3K27 mono-methylation. They showed that ES cells deficient in EED also lacked EZH2 and had significantly lower H3K27me2 and H3K27me3 levels compared to the controls. However, there was only a moderate reduction in the H3K27me1 levels [128]. Montgomery *et al.* (2005) [129] studied the murine EED, where *EEDnull/null* cells not only had reduced levels of H3K27me2 and H3K27me3, they had significantly lower H3K27me1 levels. This decreased H3K27me1 level was rescued in *EEDnull/null* ES lines when they stably expressed low levels of EED, without causing a rescue of EZH2 protein levels [129]. This suggests that EZH2 methyltransferase is not required for H3K27 mono-methylation when EED is expressed. Either EED mono-methylates H3K27 itself (Fig. 16B) or it recruits another methyltransferase which methylates H3K27. Moreover, *Suz12*^{-/-} embryos, which lack Suz12 and EZH2 but not EED, have significantly reduced levels of H3K27me2 and H3K27me3 levels. In these embryos, the levels of H3K27me1 remain unaffected [130]. This further validates that EZH2 is not essential for H3K27 mono-methylation. It is possible that a distinct complex which contains or interacts with EED mediates mono-methylation of H3K27, and subsequently passes on the role of di- and tri-methylation to EZH2.

Recent studies have shown G9a and GLP to interact with the PRC2 complex *in vivo* [125]. A genome-wide study done by Mozzetta *et al.* (2014) [125] demonstrated a significant overlap between G9a-bound genes and PRC2 target genes. Their results

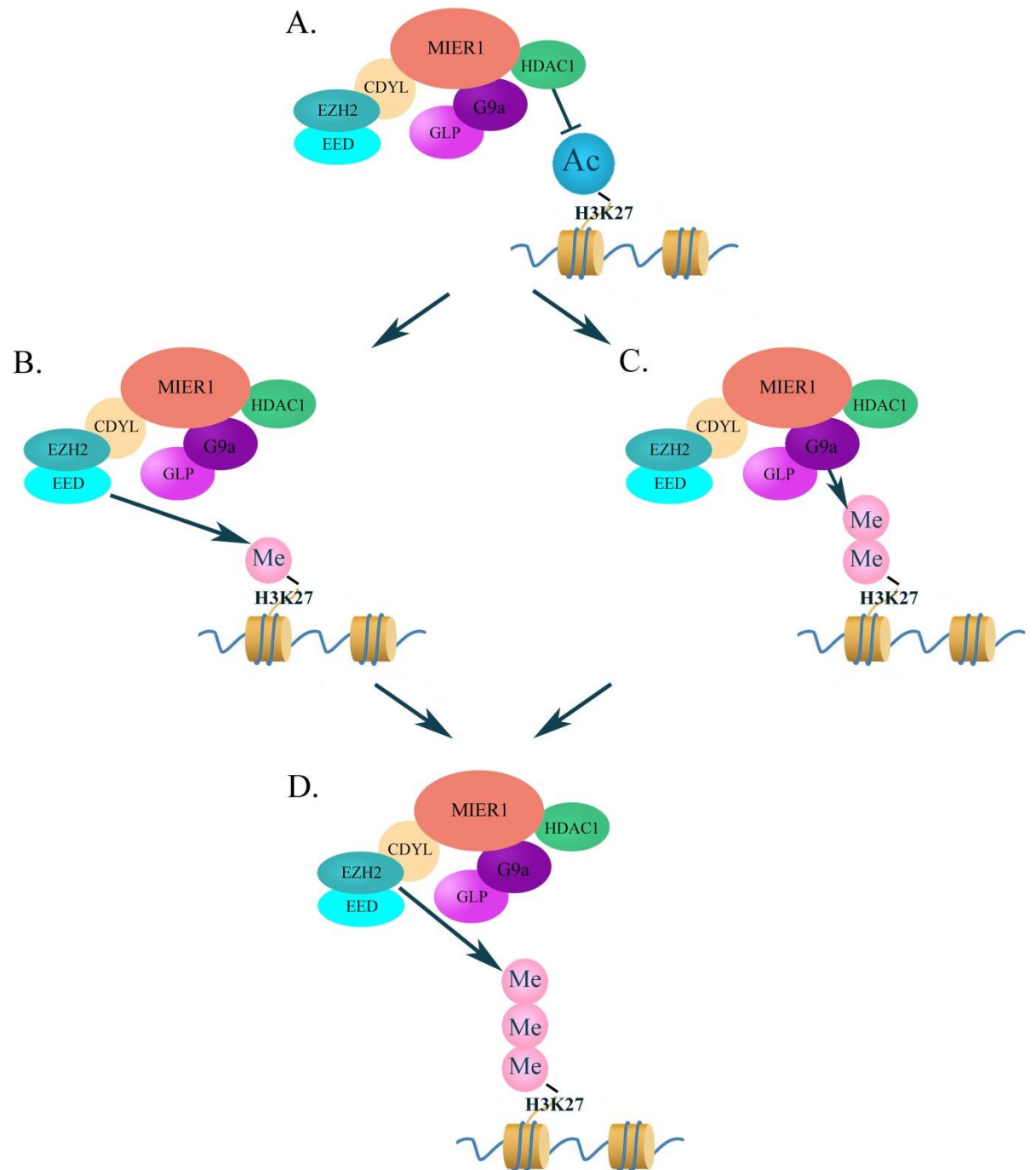
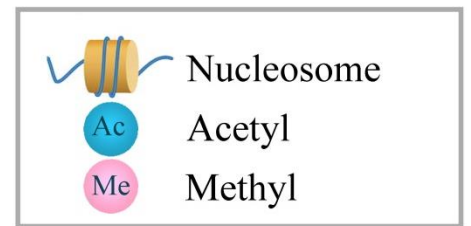
suggest G9a/GLP to cooperate with PRC2 in the regulation of gene expression, especially in regulation of a subset of developmental genes involved in neuronal differentiation. Moreover, G9a-null, murine embryonic stem cells (mESCs) had impaired EZH2 recruitment to target genes and H3K27 tri-methylation [125]. This was restored with the re-expression of wild-type G9a, but not the inactive form. This suggests that catalytically active G9a plays a role in EZH2 recruitment and H3K27 tri-methylation.

In addition to histone 3, G9a also methylates CDYL [40], and CDYL has been shown to interact with PRC2 complex [96]. It is possible that the methylation of CDYL enhances its binding to the PRC2 complex. Furthermore, MIER1 could be the mediator between these two proteins, enabling G9a to methylate CDYL, which might serve to recruit PRC2 complex. MIER1 could be the linker protein between G9a/GLP and PRC2 complexes, regulating the interplay between H3K9 and H3K27 methylation. The results in this study demonstrate the MIER1 complex to have methyltransferase activity towards H3K27, and it interacts with EZH2 via CDYL. Therefore, it is possible that MIER1 recruits G9a and its binding partners such as GLP, and together they are active at methylating H3K27 (Fig. 16C).

Another possibility is that MIER1 could be recruiting another HKMT which is responsible for mono- and di-methylation of H3K27. MIER1 then interacts with CDYL which recruits EZH2 to tri-methylate the lysine residue. The role of MIER1 in H3K27 methylation could be that of a regulator, removing acetyl group from lysine 27 and enabling mono- and di-methylation by a putative HKMT, and subsequently leading to tri-methylation by EZH2 (Fig. 16D) and repression of PRC2 regulated genes.

Figure 16: Proposed role of MIER1 in H3K27 methylation

MIER1 is a potential regulator of H3K27 methylation, removing acetyl group from lysine 27 and enabling methylation, leading to subsequent repression of PRC2 regulated genes. **A.** MIER1 recognizes and binds to acetylated H3K27 (Paterno, unpublished data). MIER1 recruits HDAC1 to acetylated H3K27 and enables it to remove the acetyl group from H3K27. By removing the acetyl group from the lysine residue, the lysine is now available to be methylated by HKMTs. **B.** EED, a PRC2 member has been shown to play a role in H3K27 methylation. It is possible for EED to initiate H3K27 methylation by mono-methylating the lysine by itself or by recruiting another HKMT. **C.** There is a possibility that G9a-GLP complex di-methylates H3K27 and enables H3K27 tri-methylation by EZH2. G9a and GLP interact with the PRC2 complex *in vivo* [128], and it is possible that after H3K27 di-methylation, G9a and GLP serve to recruit EZH2 via its interacting partners. G9a interacts with MIER1, which interacts with CDYL, as shown by this study. CDYL interacts with EZH2 [96] and could recruit it to H3K27. **D.** After H3K27 mono- and di-methylation, EZH2 is able to tri-methylate H3K27.



4.3 Implications

Tri-methylation of H3K27 is established by the PRC2 complex and it plays an important role in development. The components of the PRC2 complex, EZH2 and EED are important in the proliferation of both transformed and non-transformed human cells. These two proteins are tightly regulated by the pRB-E2F pathway [131], and deregulation of this pathway could lead to tumourgenesis. In accordance with this, EZH2 levels are significantly augmented in a large set of human tumours [131]. EZH2 overexpression in prostate and breast cancers promotes tumourgenesis and plays a role in invasiveness and metastasis [131-133]. Currently, many studies are focusing on development of specific EZH2 inhibitors to block EZH2 expression or activity. Blocking the downstream signaling cascade of EZH2 is a promising strategy for novel anticancer treatment.

According to my study, MIER1 interacts with CDYL and EZH2, and MIER1 complex immunoprecipitated from HEK-293 cells has methylation activity towards H3K27. The HKMT behind the mono- and di-methylation of this residue still remains to be verified. However, MIER1 could be the complex responsible for the initiation of H3K27 methylation, as it does recruit HDAC, which removes the acetyl group from the lysine residue. Only when K27 is not acetylated can it be methylated. Thus, in addition to developing inhibitors to alleviate EZH2 expression or activity, MIER1 could be a potential target for anticancer treatments. Inhibitors could be developed to diminish MIER1's ability to recognize the acetylated lysine, and thus lowering the capability of HDAC to deacetylate the lysine. This will lower the availability of unmodified H3K27, and thereby will lower H3K27 methylation. The role of MIER1 in the recruitment of

EZH2 should be further studied. If studies show MIER1 to be an important player in the recruitment of EZH2, then inhibitors could be designed to target the MIER1-CDYL-EZH2 interaction. This lowers the accessibility of EZH2 to its target, and subsequently will lower H3K27 methylation and repression of the PRC2 target genes.

4.4 Future studies

This study examined the interaction between MIER1, G9a, CDYL and EZH2 in an attempt to elucidate insights on the mechanism underlying H3K27 methylation. While much has been discovered regarding the interaction between these proteins, extensive study is needed to determine the role they play in H3K27 methylation. Listed below are few experiments that should be completed to further understand the mechanisms behind H3K27 methylation.

4.4.1 Determine *in vivo* interaction between endogenous MIER1 α , G9a and CDYL

This study demonstrated the *in vivo* interaction between MIER1 α and G9a, MIER1 α and CDYL, and MIER1 α , G9a and CDYL by transiently transfecting HEK-293 cells with MIER1 α , myc-tagged G9a, and flag-tagged CDYL. Subsequently, the proteins were immunoprecipitated and western blot analysis was performed. Even though this demonstrated the two proteins interacting *in vivo*, this required overexpression of both proteins in the cell. A better approach in verifying this interaction would be to perform immunoprecipitation assays using cells which endogenously express high levels of these proteins.

4.4.2 Demonstrate *in vitro* interaction between MIER1 α and CDYL

The interaction between MIER1 α and CDYL *in vivo* was demonstrated in this study, but *in vitro* interaction remains to be confirmed. Performing *in vitro* assays such as a GST pull-down assay can determine if the interaction between the two proteins is direct or requires a bridging protein. If GST pull-down assays confirm the interaction to be direct, additional *in vitro* assays can be performed to determine which domain of MIER1 and CDYL are involved in the interaction.

4.4.3 Determine whether MIER1 interacts with other members of PRC2 complex

According to this study, MIER1 α interacts with the methyltransferase, EZH2, which is a component of PRC2 complex, via CDYL. It would be interesting to determine whether MIER1 could interact with other members of the PRC2 complex such as EED and/or Suv12. If MIER1 does interact with the other PRC2 members *in vivo*, it could further elucidate the role of MIER1 in H3K27 methylation.

4.4.4 Determine the methylation activity of full length G9a

In this study we analysed the methylation activity of G9a protein which contained only its SET domain with the adjacent pre- and post-SET regions. Even though this SET domain was catalytically active towards H3K9, it did not have detectable activity towards H3K27. This gives the indication that H3K27 might not be a genuine substrate for methylation by G9a. However, to confirm this, the use of full length G9a protein is essential, as there is a possibility that other regions and domains of the protein are required for H3K27 methylation. Full-length, GST-tagged G9a could be used in the

methylation assays performed as we already have the GST-tagged G9a expressing plasmid available.

4.4.5 Determine the methylation activity of MIER1 complex

In this study, MIER1 α was immunoprecipitated from HEK-293 cells and used for *in vitro* methylation assay using ELISA. However, the inhibitors were not effective at the concentrations used. For future studies, instead of using histone N-terminal peptides [H3(21-44)] as substrate, recombinant histones or nucleosomes should be used to check MIER1 α complex methylation activity. Currently, the assay is being optimized. Once, it is established with proper controls, the methylation potential of MIER1 α complex with and without CDYL should be compared, as CDYL recruits the H3K27 methyltransferase, EZH2. This would give insights on the role of MIER1-CDYL interaction in H3K27 methylation by EZH2.

Bibliography

1. WADDINGTON, C.H., *Organisers and Genes*. 1940, Cambridge: Cambridge University Press.
2. WADDINGTON, C.H., *The Strategy of the Genes* 1957, London: Allen & Unwin.
3. CH, W., *The epigenotype* 1942. **1**: p. 18-20.
4. Das, P.M. and R. Singal, *DNA methylation and cancer*. J Clin Oncol, 2004. **22**(22): p. 4632-42.
5. Strachan T, R.A., *Human Molecular Genetics*. Cancer genetics. Vol. 2nd edition. 1999, New York: Wiley-Liss.
6. Holliday, R., *A new theory of carcinogenesis*. Br J Cancer, 1979. **40**(4): p. 513-22.
7. Dobosy, J.R., et al., *The expanding role of epigenetics in the development, diagnosis and treatment of prostate cancer and benign prostatic hyperplasia*. J Urol, 2007. **177**(3): p. 822-31.
8. Li, K.K., et al., *Chemical and biochemical approaches in the study of histone methylation and demethylation*. Med Res Rev, 2012. **32**(4): p. 815-67.
9. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications*. Nature, 2000. **403**(6765): p. 41-5.
10. Jenuwein, T. and C.D. Allis, *Translating the histone code*. Science, 2001. **293**(5532): p. 1074-80.
11. Fish, P.V., et al., *Identification of a chemical probe for bromo and extra C-terminal bromodomain inhibition through optimization of a fragment-derived hit*. J Med Chem, 2012. **55**(22): p. 9831-7.
12. Cosgrove, M.S., *Writers and readers: deconvoluting the harmonic complexity of the histone code*. Nat Struct Mol Biol, 2012. **19**(8): p. 739-40.
13. Jakovcevski, M. and S. Akbarian, *Epigenetic mechanisms in neurological disease*. Nat Med, 2012. **18**(8): p. 1194-204.
14. Taverna, S.D., et al., *How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers*. Nat Struct Mol Biol, 2007. **14**(11): p. 1025-40.

15. Eissenberg, J.C., *Structural biology of the chromodomain: form and function*. Gene, 2012. **496**(2): p. 69-78.
16. Murray, K., *The occurrence of epsilon-N-methyl lysine in histones*. Biochemistry. Vol. 3. 1964.
17. Vermeulen, M., et al., *Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers*. Cell, 2010. **142**(6): p. 967-80.
18. Payer, B. and J.T. Lee, *X chromosome dosage compensation: how mammals keep the balance*. Annu Rev Genet, 2008. **42**: p. 733-72.
19. Plath, K., et al., *Role of histone H3 lysine 27 methylation in X inactivation*. Science, 2003. **300**(5616): p. 131-5.
20. Sha, K., *A mechanistic view of genomic imprinting*. Annu Rev Genomics Hum Genet, 2008. **9**: p. 197-216.
21. Molitor, A. and W.H. Shen, *The polycomb complex PRC1: composition and function in plants*. J Genet Genomics, 2013. **40**(5): p. 231-8.
22. Lewis, E.B., *A gene complex controlling segmentation in Drosophila*. Nature, 1978. **276**(5688): p. 565-70.
23. Jürgens, G., *A group of genes controlling the spatial expression of the bithorax complex in Drosophila*. Nature, 1985. **316**: p. 153-155.
24. Margueron, R. and D. Reinberg, *The Polycomb complex PRC2 and its mark in life*. Nature, 2011. **469**(7330): p. 343-9.
25. Schwartz, Y.B. and V. Pirrotta, *Polycomb silencing mechanisms and the management of genomic programmes*. Nat Rev Genet, 2007. **8**(1): p. 9-22.
26. Allis, C.D., et al., *New nomenclature for chromatin-modifying enzymes*. Cell, 2007. **131**(4): p. 633-6.
27. Rea, S., et al., *Regulation of chromatin structure by site-specific histone H3 methyltransferases*. Nature, 2000. **406**(6796): p. 593-9.
28. Zhang, X., et al., *Structure of the Neurospora SET Domain Protein DIM-5, a Histone H3 Lysine Methyltransferase*. Cell, 2002. **111**(1): p. 117-127.
29. Helin, K. and D. Dhanak, *Chromatin proteins and modifications as drug targets*. Nature, 2013. **502**(7472): p. 480-8.

30. Schuettengruber, B., et al., *Trithorax group proteins: switching genes on and keeping them active*. Nat Rev Mol Cell Biol, 2011. **12**(12): p. 799-814.
31. Greer, E.L. and Y. Shi, *Histone methylation: a dynamic mark in health, disease and inheritance*. Nat Rev Genet, 2012. **13**(5): p. 343-57.
32. Jia, H., et al., *Genome-wide computational identification and manual annotation of human long noncoding RNA genes*. Rna, 2010. **16**(8): p. 1478-87.
33. Zilberman, D., X. Cao, and S.E. Jacobsen, *ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation*. Science, 2003. **299**(5607): p. 716-9.
34. Fukagawa, T., et al., *Dicer is essential for formation of the heterochromatin structure in vertebrate cells*. Nat Cell Biol, 2004. **6**(8): p. 784-91.
35. Ogawa, Y., B.K. Sun, and J.T. Lee, *Intersection of the RNA interference and X-inactivation pathways*. Science, 2008. **320**(5881): p. 1336-41.
36. Rajakumara, E., et al., *A dual flip-out mechanism for 5mC recognition by the Arabidopsis SUVH5 SRA domain and its impact on DNA methylation and H3K9 dimethylation in vivo*. Genes Dev, 2011. **25**(2): p. 137-52.
37. Johnson, L.M., et al., *The SRA methyl-cytosine-binding domain links DNA and histone methylation*. Curr Biol, 2007. **17**(4): p. 379-84.
38. Tachibana, M., et al., *Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3*. J Biol Chem, 2001. **276**(27): p. 25309-17.
39. Patnaik, D., et al., *Substrate specificity and kinetic mechanism of mammalian G9a histone H3 methyltransferase*. J Biol Chem, 2004. **279**(51): p. 53248-58.
40. Sampath, S.C., et al., *Methylation of a histone mimic within the histone methyltransferase G9a regulates protein complex assembly*. Mol Cell, 2007. **27**(4): p. 596-608.
41. Huang, J., et al., *G9a and Glp methylate lysine 373 in the tumor suppressor p53*. J Biol Chem, 2010. **285**(13): p. 9636-41.
42. Rathert, P., et al., *Protein lysine methyltransferase G9a acts on non-histone targets*. Nat Chem Biol, 2008. **4**(6): p. 344-6.
43. Lee, J.S., et al., *Negative regulation of hypoxic responses via induced Reptin methylation*. Mol Cell, 2010. **39**(1): p. 71-85.

44. Chin, H.G., et al., *Automethylation of G9a and its implication in wider substrate specificity and HP1 binding*. Nucleic Acids Res, 2007. **35**(21): p. 7313-23.
45. Collins, R. and X. Cheng, *A case study in cross-talk: the histone lysine methyltransferases G9a and GLP*. Nucleic Acids Res, 2010. **38**(11): p. 3503-11.
46. Tachibana, M., et al., *Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9*. Genes Dev, 2005. **19**(7): p. 815-26.
47. Milner, C.M. and R.D. Campbell, *The G9a gene in the human major histocompatibility complex encodes a novel protein containing ankyrin-like repeats*. Biochem J, 1993. **290** (Pt 3): p. 811-8.
48. Bourbon, H.M., B. Lapeyre, and F. Amalric, *Structure of the mouse nucleolin gene. The complete sequence reveals that each RNA binding domain is encoded by two independent exons*. J Mol Biol, 1988. **200**(4): p. 627-38.
49. Wen, L., et al., *A human placental cDNA clone that encodes nonhistone chromosomal protein HMG-I*. Nucleic Acids Res, 1989. **17**(3): p. 1197-214.
50. Earnshaw, W.C., *Anionic regions in nuclear proteins*. J Cell Biol, 1987. **105**(4): p. 1479-82.
51. Earnshaw, W.C., et al., *Molecular cloning of cDNA for CENP-B, the major human centromere autoantigen*. J Cell Biol, 1987. **104**(4): p. 817-29.
52. Roopra, A., et al., *Localized domains of G9a-mediated histone methylation are required for silencing of neuronal genes*. Mol Cell, 2004. **14**(6): p. 727-38.
53. Collins, R.E., et al., *The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules*. Nat Struct Mol Biol, 2008. **15**(3): p. 245-50.
54. Dillon, S.C., et al., *The SET-domain protein superfamily: protein lysine methyltransferases*. Genome Biol, 2005. **6**(8): p. 227.
55. Taylor, W.R., et al., *A knot or not a knot? SETting the record 'straight' on proteins*. Comput Biol Chem, 2003. **27**(1): p. 11-5.
56. Trievel, R.C., et al., *Mechanism of multiple lysine methylation by the SET domain enzyme Rubisco LSM1*. Nat Struct Biol, 2003. **10**(7): p. 545-52.
57. Krishnan, S., S. Horowitz, and R.C. Trievel, *Structure and function of histone H3 lysine 9 methyltransferases and demethylases*. Chembiochem, 2011. **12**(2): p. 254-63.

58. Shahbazian, M.D., K. Zhang, and M. Grunstein, *Histone H2B ubiquitylation controls processive methylation but not monomethylation by Dot1 and Set1*. Mol Cell, 2005. **19**(2): p. 271-7.
59. Chin, H.G., et al., *Sequence specificity and role of proximal amino acids of the histone H3 tail on catalysis of murine G9A lysine 9 histone H3 methyltransferase*. Biochemistry, 2005. **44**(39): p. 12998-3006.
60. Sancho, M., et al., *Depletion of human histone H1 variants uncovers specific roles in gene expression and cell growth*. PLoS Genet, 2008. **4**(10): p. e1000227.
61. Pless, O., et al., *G9a-mediated lysine methylation alters the function of CCAAT/enhancer-binding protein-beta*. J Biol Chem, 2008. **283**(39): p. 26357-63.
62. Happel, N. and D. Doenecke, *Histone H1 and its isoforms: contribution to chromatin structure and function*. Gene, 2009. **431**(1-2): p. 1-12.
63. Daujat, S., et al., *HP1 binds specifically to Lys26-methylated histone H1.4, whereas simultaneous Ser27 phosphorylation blocks HP1 binding*. J Biol Chem, 2005. **280**(45): p. 38090-5.
64. Trojer, P. and D. Reinberg, *Beyond histone methyl-lysine binding: how malignant brain tumor (MBT) protein L3MBTL1 impacts chromatin structure*. Cell Cycle, 2008. **7**(5): p. 578-85.
65. Lee, H., R. Habas, and C. Abate-Shen, *MSX1 cooperates with histone H1b for inhibition of transcription and myogenesis*. Science, 2004. **304**(5677): p. 1675-8.
66. Feldman, N., et al., *G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis*. Nat Cell Biol, 2006. **8**(2): p. 188-94.
67. El Gazzar, M., et al., *G9a and HP1 couple histone and DNA methylation to TNFalpha transcription silencing during endotoxin tolerance*. J Biol Chem, 2008. **283**(47): p. 32198-208.
68. Grewal, S.I. and D. Moazed, *Heterochromatin and epigenetic control of gene expression*. Science, 2003. **301**(5634): p. 798-802.
69. Kwon, S.H. and J.L. Workman, *The heterochromatin protein 1 (HP1) family: put away a bias toward HP1*. Mol Cells, 2008. **26**(3): p. 217-27.
70. Epsztejn-Litman, S., et al., *De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes*. Nat Struct Mol Biol, 2008. **15**(11): p. 1176-83.

71. Esteve, P.O., et al., *Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication*. Genes Dev, 2006. **20**(22): p. 3089-103.
72. Tachibana, M., et al., *G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis*. Genes Dev, 2002. **16**(14): p. 1779-91.
73. Shinkai, Y. and M. Tachibana, *H3K9 methyltransferase G9a and the related molecule GLP*. Genes Dev, 2011. **25**(8): p. 781-8.
74. Tachibana, M., et al., *Functional dynamics of H3K9 methylation during meiotic prophase progression*. Embo j, 2007. **26**(14): p. 3346-59.
75. Thomas, L.R., et al., *Functional analysis of histone methyltransferase g9a in B and T lymphocytes*. J Immunol, 2008. **181**(1): p. 485-93.
76. Lehnertz, B., et al., *Activating and inhibitory functions for the histone lysine methyltransferase G9a in T helper cell differentiation and function*. J Exp Med, 2010. **207**(5): p. 915-22.
77. Schaefer, A., et al., *Control of cognition and adaptive behavior by the GLP/G9a epigenetic suppressor complex*. Neuron, 2009. **64**(5): p. 678-91.
78. Kramer, J.M., et al., *Epigenetic regulation of learning and memory by Drosophila EHMT/G9a*. PLoS Biol, 2011. **9**(1): p. e1000569.
79. Maze, I., et al., *Essential role of the histone methyltransferase G9a in cocaine-induced plasticity*. Science, 2010. **327**(5962): p. 213-6.
80. Gyory, I., et al., *PRDI-BF1 recruits the histone H3 methyltransferase G9a in transcriptional silencing*. Nat Immunol, 2004. **5**(3): p. 299-308.
81. Nishio, H. and M.J. Walsh, *CCAAT displacement protein/cut homolog recruits G9a histone lysine methyltransferase to repress transcription*. Proc Natl Acad Sci U S A, 2004. **101**(31): p. 11257-62.
82. Cho, H.S., et al., *Enhanced expression of EHMT2 is involved in the proliferation of cancer cells through negative regulation of SIAH1*. Neoplasia, 2011. **13**(8): p. 676-84.
83. Chen, M.W., et al., *H3K9 histone methyltransferase G9a promotes lung cancer invasion and metastasis by silencing the cell adhesion molecule Ep-CAM*. Cancer Res, 2010. **70**(20): p. 7830-40.

84. Dong, C., et al., *G9a interacts with Snail and is critical for Snail-mediated E-cadherin repression in human breast cancer*. J Clin Invest, 2012. **122**(4): p. 1469-86.
85. Shi, Y., et al., *Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds*. Cell Stem Cell, 2008. **3**(5): p. 568-74.
86. Liu, F., et al., *Discovery of a 2,4-diamino-7-aminoalkoxyquinazoline as a potent and selective inhibitor of histone lysine methyltransferase G9a*. J Med Chem, 2009. **52**(24): p. 7950-3.
87. Chang, Y., et al., *Adding a lysine mimic in the design of potent inhibitors of histone lysine methyltransferases*. J Mol Biol, 2010. **400**(1): p. 1-7.
88. Lahn, B.T. and D.C. Page, *Functional coherence of the human Y chromosome*. Science, 1997. **278**(5338): p. 675-80.
89. Lahn, B.T. and D.C. Page, *Retroposition of autosomal mRNA yielded testis-specific gene family on human Y chromosome*. Nat Genet, 1999. **21**(4): p. 429-33.
90. Caron, C., et al., *Cdyl: a new transcriptional co-repressor*. EMBO Rep, 2003. **4**(9): p. 877-82.
91. Lahn, B.T., et al., *Previously uncharacterized histone acetyltransferases implicated in mammalian spermatogenesis*. Proc Natl Acad Sci U S A, 2002. **99**(13): p. 8707-12.
92. Shi, Y., et al., *Coordinated histone modifications mediated by a CtBP co-repressor complex*. Nature, 2003. **422**(6933): p. 735-8.
93. Mulligan, P., et al., *CDYL bridges REST and histone methyltransferases for gene repression and suppression of cellular transformation*. Mol Cell, 2008. **32**(5): p. 718-26.
94. Kuppaswamy, M., et al., *Role of the PLDLS-binding cleft region of CtBP1 in recruitment of core and auxiliary components of the corepressor complex*. Mol Cell Biol, 2008. **28**(1): p. 269-81.
95. Li, X., et al., *Functional consequences of new exon acquisition in mammalian chromodomain Y-like (CDYL) genes*. Trends Genet, 2007. **23**(9): p. 427-31.
96. Zhang, Y., et al., *Corepressor protein CDYL functions as a molecular bridge between polycomb repressor complex 2 and repressive chromatin mark trimethylated histone lysine 27*. J Biol Chem, 2011. **286**(49): p. 42414-25.

97. Mark, M., F.M. Rijli, and P. Chambon, *Homeobox genes in embryogenesis and pathogenesis*. *Pediatr Res*, 1997. **42**(4): p. 421-9.
98. Cao, R., et al., *Role of histone H3 lysine 27 methylation in Polycomb-group silencing*. *Science*, 2002. **298**(5595): p. 1039-43.
99. Paterno, G.D., et al., *cDNA cloning of a novel, developmentally regulated immediate early gene activated by fibroblast growth factor and encoding a nuclear protein*. *J Biol Chem*, 1997. **272**(41): p. 25591-5.
100. Paterno, G.D., et al., *Molecular cloning of human er1 cDNA and its differential expression in breast tumours and tumour-derived cell lines*. *Gene*, 1998. **222**(1): p. 77-82.
101. Paterno, G.D., et al., *Genomic organization of the human mi-er1 gene and characterization of alternatively spliced isoforms: regulated use of a facultative intron determines subcellular localization*. *Gene*, 2002. **295**(1): p. 79-88.
102. McCarthy, P.L., et al., *Changes in subcellular localisation of MI-ER1 alpha, a novel oestrogen receptor-alpha interacting protein, is associated with breast cancer progression*. *Br J Cancer*, 2008. **99**(4): p. 639-46.
103. Clements, J.A., et al., *Differential splicing alters subcellular localization of the alpha but not beta isoform of the MIER1 transcriptional regulator in breast cancer cells*. *PLoS One*, 2012. **7**(2): p. e32499.
104. Heery, D.M., et al., *A signature motif in transcriptional co-activators mediates binding to nuclear receptors*. *Nature*, 1997. **387**(6634): p. 733-6.
105. Solari, F., A. Bateman, and J. Ahringer, *The Caenorhabditis elegans genes egl-27 and egr-1 are similar to MTA1, a member of a chromatin regulatory complex, and are redundantly required for embryonic patterning*. *Development*, 1999. **126**(11): p. 2483-94.
106. Ding, Z., L.L. Gillespie, and G.D. Paterno, *Human MI-ER1 alpha and beta function as transcriptional repressors by recruitment of histone deacetylase 1 to their conserved ELM2 domain*. *Mol Cell Biol*, 2003. **23**(1): p. 250-8.
107. Blackmore, T.M., et al., *The transcriptional cofactor MIER1-beta negatively regulates histone acetyltransferase activity of the CREB-binding protein*. *BMC Res Notes*, 2008. **1**: p. 68.
108. Aasland, R., A.F. Stewart, and T. Gibson, *The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB*. *Trends Biochem Sci*, 1996. **21**(3): p. 87-8.

109. Ordentlich, P., et al., *Unique forms of human and mouse nuclear receptor corepressor SMRT*. Proc Natl Acad Sci U S A, 1999. **96**(6): p. 2639-44.
110. Toh, Y., et al., *Molecular analysis of a candidate metastasis-associated gene, MTA1: possible interaction with histone deacetylase 1*. J Exp Clin Cancer Res, 2000. **19**(1): p. 105-11.
111. Wade, P.A., et al., *Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation*. Nat Genet, 1999. **23**(1): p. 62-6.
112. You, A., et al., *CoREST is an integral component of the CoREST- human histone deacetylase complex*. Proc Natl Acad Sci U S A, 2001. **98**(4): p. 1454-8.
113. Guenther, M.G., O. Barak, and M.A. Lazar, *The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3*. Mol Cell Biol, 2001. **21**(18): p. 6091-101.
114. Sterner, D.E., et al., *The SANT domain of Ada2 is required for normal acetylation of histones by the yeast SAGA complex*. J Biol Chem, 2002. **277**(10): p. 8178-86.
115. Ding, Z., et al., *The SANT domain of human MI-ER1 interacts with Sp1 to interfere with GC box recognition and repress transcription from its own promoter*. J Biol Chem, 2004. **279**(27): p. 28009-16.
116. Wang, L., et al., *Atrophin recruits HDAC1/2 and G9a to modify histone H3K9 and to determine cell fates*. EMBO Rep, 2008. **9**(6): p. 555-62.
117. Miller, A.J., *Le Châtelier's principle and the equilibrium constant*. Journal of Chemical Education, 1954. **31**(9): p. 455.
118. Thorne, L.B., et al., *Cloning and characterization of the mouse ortholog of mi-er1*. DNA Seq, 2005. **16**(3): p. 237-40.
119. Boyer, L.A., R.R. Latek, and C.L. Peterson, *The SANT domain: a unique histone-tail-binding module?* Nat Rev Mol Cell Biol, 2004. **5**(2): p. 158-63.
120. Margueron, R., et al., *Role of the polycomb protein EED in the propagation of repressive histone marks*. Nature, 2009. **461**(7265): p. 762-7.
121. Xu, C., et al., *Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2)*. Proc Natl Acad Sci U S A, 2010. **107**(45): p. 19266-71.
122. Bantscheff, M., et al., *Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes*. Nat Biotechnol, 2011. **29**(3): p. 255-65.

123. Czermin, B., et al., *Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites*. Cell, 2002. **111**(2): p. 185-96.
124. Muller, J., et al., *Histone methyltransferase activity of a Drosophila Polycomb group repressor complex*. Cell, 2002. **111**(2): p. 197-208.
125. Mozzetta, C., et al., *The Histone H3 Lysine 9 Methyltransferases G9a and GLP Regulate Polycomb Repressive Complex 2-Mediated Gene Silencing*. Mol Cell, 2014. **53**(2): p. 277-89.
126. Li, Y., et al., *The target of the NSD family of histone lysine methyltransferases depends on the nature of the substrate*. J Biol Chem, 2009. **284**(49): p. 34283-95.
127. Singer, M.S., et al., *Identification of high-copy disruptors of telomeric silencing in Saccharomyces cerevisiae*. Genetics, 1998. **150**(2): p. 613-32.
128. Schoeftner, S., et al., *Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing*. Embo j, 2006. **25**(13): p. 3110-22.
129. Montgomery, N.D., et al., *The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation*. Curr Biol, 2005. **15**(10): p. 942-7.
130. Pasini, D., et al., *Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity*. The EMBO Journal, 2004. **23**(20): p. 4061-4071.
131. Bracken, A.P., et al., *EZH2 is downstream of the pRB-E2F pathway, essential for proliferation and amplified in cancer*. Embo j, 2003. **22**(20): p. 5323-35.
132. Varambally, S., et al., *The polycomb group protein EZH2 is involved in progression of prostate cancer*. Nature, 2002. **419**(6907): p. 624-9.
133. Kleer, C.G., et al., *EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells*. Proc Natl Acad Sci U S A, 2003. **100**(20): p. 11606-11.